5-1-1977

Punctuation Frequency in Selected American Daily Newspapers

Andrew Malcolm

PUNCTUATION FREQUENCY IN

SELECTED AMERICAN DAILY NEWSPAPERS

By

Andrew Malcolm

A thesis submitted in partial fulfillment of the
requirements for the degree of Master of Science in the
School of Printing in the College of Graphic Arts and Photography
of the Rochester Institute of Technology

May 1977

Thesis advisor: Mrs. Ruth Terry

Certificate of Approval--Master's Thesis

School of Printing
Rochester Institute of Technology
Rochester, New York

CERTIFICATE OF APPROVAL

MASTER'S THESIS

This is to certify that the Master's Thesis of

Andrew Malcolm

with a major in Printing Technology has been approved by the Thesis Committee as satisfactory for the thesis requirement for the Master of Science degree at the convocation of May 21, 1977.

Thesis Committee:

Ruth L. Terry
Ruth L. Terry
Thesis Advisor

Julius Silver
Julius L. Silver, Ph.D.
Graduate Advisor

Mark F. Guldin
Mark F. Guldin, Ph D.
Director or Designate

ACKNOWLEDGEMENTS

I, Andrew Malcolm, am pleased to acknowledge assistance of Mrs. Ruth Terry, thesis advisor, member of the adjunct faculty of The School of Printing, and long-time friend. Her suggestions both in the area of the design of this experiment and stylistic features of its presentation are appreciated and acknowledged. I am also pleased to acknowledge the tireless efforts of the staff of the Wallace Memorial Library in obtaining newspaper samples, including Ms. Janice Linehan, Mrs. Gladys Taylor, Ms. MaryAnn Young, and others unknown to me. I am indebted to Ms. Mary Lynn Vickers who helped search for relevant literature. I thank Professor Robert S. Zais of Kent State University for granting permission to reprint a portion of his 1950 master's thesis as Appendix A to this thesis. I am indebted to Dr. Kathleen E. Crandall and the National Technical Institute for the Deaf for support and assistance throughout the preparation of this thesis.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

LIST OF TABLES

LIST OF TABLES (CONTINUED)

LIST OF TABLES (CONTINUED)

LIST OF FIGURES

ABSTRACT

American daily newspapers present millions of punctuated sentences to the public, yet--after an extensive search of literature for similar studies--no comprehensive study of punctuation frequency seems to have been made before this study.

It is hypothesized (1) that punctuation appears at predictable rates in the main news section of American daily newspapers, (2) that statistically significant differences exist in punctuation frequency between each of five selected newspapers and--by group--between those of large and small cities, and (3) that statistically significant differences exist in punctuation frequency between 1976, 1966, 1956 and 1946.

Five newspapers were selected for the study: The New York Times, The Los Angeles Times, The Chicago Tribune, The Atlanta Journal and Constitution, and The Denver Post. The sample dates are February 29, 1976--the first Sunday following the approval of this study--, and the three decades preceeding it, viz., February 27, 1966, February 26, 1956, and March 3, 1946. Samples included the first page of the main news section of each paper. After preliminary studies to determine an appropriate sample size, it was decided to use all words appearing on the first page of each of the twenty newspapers.

Twelve marks of punctuation and twenty-three subcategories of use were enumerated as they occurred for each of the 52,184 words which appear on the first page of the twenty newspapers. These include two uses of the apostrophe, colons, eight uses of the comma, dashes, ellipses, terminal and other use of the exclamation point, three uses of the hyphen, parentheses, three uses of the period, terminal and other use of the question mark, two uses of the quotation mark, and three uses of the semicolon.

Frequencies were found within all twenty cells which ranged from no occurrences to 111 commas per kiloword. Ninety-five percent confidence intervals were calculated for each use in each cell. These ranged from $\pm$ 12.41% to $\pm$ 196%.

Using the summary data for all punctuation for each of the cells, mean punctuation frequencies and 95% confidence intervals were calculated for (a) each cell, (b) for each newspaper, (c) for large and small cities, (d) for each decade, and (e) for all data. These means and confidence intervals were found for terminal marks and for intrasentence marks as well as all marks. It was found that all punctuation occurred at the rate of 199 marks per kiloword $\pm$ 1.9%, that there were 24 words per sentence $\pm$ 4.1%, and that intrasentence punctuation occurred at the rate of 156 marks per kiloword $\pm$ 2.3%.

Analyses of variance were conducted to determine if any of the differences across cells were statistically significant. The number of words per sentence was significantly different by both decade and newspaper at a probability level of 99%. The frequency of all punctuation, and the frequency of intrasentence punctuation did not change by either decade or newspaper.

CHAPTER 1. INTRODUCTION

I-1. Justification

American daily newspapers present millions of punctuated sentences to the public. Along with other publications, these newspapers reflect punctuation use in modern American society. Although certain variations in punctuation do not affect understanding, other variations in punctuation do. It is necessary to have a common referent for words for people to communicate. Although it is of lesser importance, it is also necessary to have a common meaning for punctuation to insure precise written communication between people. Some writers follow the rules given in one or more of the many guides that are available. Other writers punctuate intuitively. It is widely believed that significant variations in punctuation occur.

Persons engaged in writing news text could improve the probability of being precisely understood by their readers if they used punctuation in a standard manner. This is not to suggest that variations in current use are very great, but rather that variations do exist and that such variations reduce the effective communication which occurs between writer and reader. A punctuation standard based on punctuation frequency, rather than on 'rules,' established by authorities, will be useful to editors. A frequency-based standard would provide editors with justification for editors' changes.

Many persons believe that language is the most complex of all human behaviors. Writing is one of the two major modes of expressive language. Writing behavior should be of interest to anyone involved in any aspect of the publishing industry because the vast majority of publishing includes words that require punctuation. Although not of primary interest to all persons who are interested in language, punctuation is an important aspect of writing behavior.

I-1.1 Application to Composing Machine Design

Designers of composing machines need to know the alphabetic and punctuation frequency for the language to be composed on their machines. The mechanical designer of a new photocomposition machine needs to know the frequency of use of each type image in his machine to insure a reliable design. More wear will be associated with frequently used letters and marks of punctuation than with infrequently used ones. In strike-on systems, this must be balanced against the striking force and the surface area of the image. Punctuation presents a different problem than letters because the most frequently used marks of punctuation are also the marks with the smallest surface areas.

I-1.2 Application to NTID

The National Technical Institute for the Deaf (NTID) is training students who, in the large majority, have anomalous writing behavior. Donald Johnson (1976) reported that these anomalies are traceable to

having become deaf prior to learning the auditory language of their environment.[1] For several years, NTID has been developing methods and means to present written language to deaf students in a manner which more closely approaches spoken language. Punctuation is a representation of various prosodic features of spoken English. The comma, the semicolon, the dash, the colon, the period, and the paragraph are representative of increasingly longer pauses in natural speech. The exclamation point and question mark are representative of speech inflections. Obviously, punctuation plays an important part in the printed representation of spoken language.

To represent speech in print as accurately as our current technology allows, NTID is using three glossynographs to present printed language to students in a manner approximating speech. A glossynograph (gloss = speech + syn = with + graph = writing) is a device which allows its operator to present words on a television screen together with speech for the deaf (q.v. Appendix G). Words may be produced directly by typing on the glossynograph keyboard, or--using unseen codes--by recalling words stored on magnetic tape. However poorly the glossynograph represents speech, it is more representative of speech than the printed page. NTID hopes to improve the representation by the development of improved equipment and programs. Although it is not the main purpose of this study, it represents an important future application of the data and is further justification. (The equipment aspects of this improvement are irrelevant to this thesis.) Students see letters which form words which, in turn, form sentences as they appear sequentially on television screens. These sentences are punctuated.

NTID has begun to conduct a study to determine, with some degree of precision, the time characteristics of speech as they relate to the graphic representation of speech. Among other factors, this will determine the time consumed, as represented by punctuation, by a typical speaker. The glossynograph uses 0.3 seconds for a comma and 1.0 seconds for a period. These times have not been established scientifically and may be easily changed on glossynographic recordings. When designing programs for the glossynograph, it will be necessary to establish standards for punctuation. The results of the proposed study will be used as the basis for the needed standard.

While designing programs for the glossynograph, it is necessary to establish a standard of writing which includes punctuation. The proposed study will provide an objective description of contemporary English punctuation as used in American newspapers.

I-1.3 Application to Newswriters and Editors

Newswriters and editors may be concerned with how the punctuation content of their manuscripts compares with similar stories in leading American daily newspapers. Other writers should be interested in this information. The data collected and analyzed in this thesis provides writers and editors with normal punctuation frequency data for comparison.

I-1.4 Prior Research on Punctuation Frequency

A search of literature on punctuation frequency was made to find similar research to that reported herein. This literature search located only one significant report, that of Harry Shaw (1963) who found that punctuation use declined between 1851 and 1961. He tallied punctuation by type from 100 consecutive sentences in copies of The New York Times. He found that the number of marks of punctuation per sentence declined from 3.39 in 1851 to 2.55 in 1962.[2] Shaw did some other studies. Having found only one prior study, first one, and later another professional library researcher were hired to locate other relevant work. No other important work on punctuation frequency was found. Ms. Mary Lynn Vickers (1975), one of the researchers, reported, "The conclusion seems to be that no comprehensive study of punctuation frequency has ever been done, ... it certainly seems that there is a dearth of relevant information available from the wide variety of sources consulted."[3]

I-2 Summary of Chapter I

It has been shown in this chapter that both practical and intellectual needs exist for the data that would be produced from the proposed study. The data should be valuable both to those in printing, and to those otherwise interested in English.

It has been shown in this chapter that the data produced by the proposed study would be useful to those designing printing composition equipment. For the reasons enumerated it appears that there is a substantial need for an in-depth study of punctuation frequency at this time.

It has been shown that the development of the glossynograph at NTID will find application for the data collected and analyzed in this thesis.

It has been shown that there is a dearth of research in this area, and that fact alone demands resolution. The only similar research found during an extensive professional literature search was that of Shaw who made no distinction between the different uses of the same mark of punctuation. The comma has 16 uses as defined by the Government Printing Office Style Guide (1967).[4] Most other marks of punctuation have multiple uses. Other facts which affect punctuation frequency were not considered by Shaw. It is obvious that the number of words per sentence, and the number of clauses (and certain types of phrases) per sentence affect punctuation frequency. Even if Shaw's data were adequate, it is fifteen years old. Shaw showed that punctutaion declined from 3.39 to 2.55 marks per sentence in 110 years. If the decline has continued at this same rate, one might expect a replication of his work to show a decline to 2.43 marks of punctuation per sentence by 1976, by extrapolation of the data.

ENDNOTES FOR CHAPTER I

[1]Donald D. Johnson, Communication Characteristics of NTID Students, Rochester, NY: Rochester Institute of Technology, National Technical Institute for the Deaf, 1976, pp. 25-29.

[2]Harry Shaw, Punctuate It Right, New York: Barnes & Noble, Inc., 1963, pp. 11-13.

[3]Mary Lynn Vickers, "Report on a Literature Search: Statistical Studies of Punctuation Frequency," 1975. (Ms. Vicker's report is included as Appendix B to this thesis.)

[4]Style Manual (abridged), Washington: United States Government Printing Office, 1967, pp. 139-141.

CHAPTER II. LITERATURE SEARCH

A search of literature on punctuation frequency was made to find similar research. This search located a text by Harry Shaw (1963) who found that punctuation use declined between 1851 and 1961. He tallied punctuation by type from 100 consecutive sentences in copies of The New York Times. He used the Times issues of September 24, 25, and 30, 1851; February 2 and 3, 1911; and February 1 and 2, 1961. Table I shows Shaw's data. He found that the number of marks of punctuation per sentence declined from 3.39 in 1851 to 2.55 in 1961. Shaw also states, ". . .that random issues . . . of five newspapers . . . show approximately the (same) decrease revealed in the New York Times."[1]

Shaw also studied magazines. He used the Atlantic Monthly issues of February 1911, and February 1961; and the Harper's Magazine issues of February 1851 and February 1961. Table II shows Shaw's data. He found no significant change in the most recent half century of Atlantic Monthly, but a century of Harper's shows a decrease in punctuation frequency.[2]

It should be noted that the total number of terminal marks of punctuation did not decline because it is these 100 marks of punctuation in each study that determine what are defined as sentences. If we eliminate the terminal marks from the statistics in the Times study, we show a greater percentage decline in intrasentence punctuation. Intrasentence punctuation per sentence in Shaw's study of the Times declined from 2.39 in 1851 to 1.55 in 1961. Intrasentence punctuation in Shaw's study of Harper's declined from 1.99 in 1851 to 1.60 in 1961.

TABLE I.

SHAW'S HISTORICAL TREND STUDY OF PUNCTUATION IN THE NEW YORK TIMES[3]

Source: New York Times	Terminal Marks of Punctuation					Intrasentence Marks of Punctuation						Total
	.	!	?	:	,	;	()	?	--	:	,--	
1851	98	1	0	1	195	28	4	1	7	3	1	339
1911	94	2	3	1	167	6	2	0	11	0	0	286
1961	99	0	1	0	125	10	3	0	17	0	0	255

TABLE II.

SHAW'S HISTORICAL TREND STUDY OF PUNCTUATION IN ATLANTIC MONTHLY AND HARPER'S[4]

Source: Magazine	Terminal Marks of Punctuation					Intrasentence Marks of Punctuation						Total
	.	!	?	:	,	;	()	!	--	:	,--	
Atlantic 1911	90	5	5	0	172	14	0	1	13	14	9	310
Atlantic 1961	97	0	0	3	195	4	5	0	4	1	0	309
Harper's 1851	99	0	0	1	175	17	0	0	7	0	0	299
Harper's 1961	90	0	8	2	118	11	6	0	24	1	0	260

Shaw did not show the number of words in his study and, without reexamining his samples, one does not know the cause of the apparent increase he found. If the number of words per sentence changed between his sample intervals, one simply cannot say that the true frequency of punctuation changed. Also, Shaw did not use any statistical procedures beyond the simple mean of occurrence.

It is not the purpose of the proposed study to consider the history of punctuation, but certainly this is a related and interesting background to investigate. Robert Stanley Zais (1950) compiled a history of punctuation in his master's thesis for Brown University. Zais asserts that, "Although men have been writing for at least six thousand years, punctuation as we know it is only a few hundred years old."

In his thesis, Zais traces the history of the major marks of punctuation from their appearance to contemporary times (1950). Because Zais' work may be of value to some readers of this thesis, and because it is not easily available, it is included as Appendix A.

Having found very little literature on punctuation frequency himself, the candidate employed the services of a professional literature search organization, Research Assistance, Inc., of Los Angeles. Their search yielded no useful data. Not being satisfied with this, the candidate employed Ms. Mary Lynn Vickers of the Rochester Regional Research Library Council.[8] Ms. Vickers' report is provided as Appendix B to this thesis. Ms. Vickers reports that,

> The conclusion seems to be that no comprehensive study of statistical frequency of punctuation has ever been done, or if it has been done, it is certainly not in a useful, widely-used or known format. For people interested in such a study, it certainly seems that there is a dearth of relevant information available from the wide variety of sources consulted.[9]

Ms. Vickers' literature research located the thesis of Zais which has been discussed. She found a number of other references which were investigated more fully by the candidate. These references led to still others, but no similar bodies of data were found.

ENDNOTES FOR CHAPTER II

[1]Harry Shaw, Punctuate It Right!, New York: Barnes & Noble, Inc., 1963, p. 12.

[2]Ibid., pp. 12, 13.

[3]Ibid., p. 12.

[4]Ibid., p. 13.

[5]Robert S. Zais, A Study to Determine Whether a Punctuation Symbol's Characteristics Influence the Effectiveness of a Method of Teaching Punctuation, unpublished Master's thesis, Department of English and Education, Brown University, 1963, pp. 1-21. (Pages 1 through 21 of Professor Zais' thesis is included as Appendix A to this thesis.)

[6]Ibid., p. 1.

[7]Ibid., p. 13.

[8]Mary Lynn Vickers, "Report on a Literature Search: Statistical Studies of Punctuation Frequency," 1975. (Ms. Vicker's report is included as Appendix B to this thesis.)

[9]Ibid., p. 3.

CHAPTER III. THEORETICAL BASIS OF STUDY

This study has determined, with a degree of statistical probability, the frequency of use of punctuation in certain major American daily newspapers by category of use.

Twelve common marks of punctuation were examined, viz., (arranged alphabetically) (1) the apostrophe ('), (2) the colon (:), (3) the comma (,), (4) the dash (--), (5) the ellipsis (. . .), (6) the exclamation point (!), (7) the hyphen (-), (8) parentheses (()), (9) the period (.), (10) the question mark (?), (11) quotation marks ("" & ''), and (12) the semicolon (;).

Each of these marks of punctuation were examined for their applications. The applications include, those identified in Appendix C. Some of the marks shown in Appendix C are restricted to a single use, while others have wide and varying use. To determine only the total number of occurrences of a mark would not be adequate, because the use of single marks differs, e.g., the hyphen is used for joining compound words, for joining adjectives in unit modifiers, for breaking words at the ends of lines, and for a few other uses.

In this thesis, the trend in frequency of the twelve marks of punctuation has been surveyed, their various uses identified, and regional differences examined.

Punctuation appears with regularity in publications. It is possible to take a sample of printed matter and calculate a mean frequency of occurrence of various marks of punctuation. It is possible with modern

statistical sampling techniques to determine, with a known probability of being correct, if the sample mean is the same as the true mean. Because different marks of punctuation appear with different frequencies, it follows that larger samples are required to obtain reliable results for certain infrequent punctuation. Also, certain uses of a given mark of punctuation appear with different frequencies. Again, judicious sampling can allow statistical inferences to be made about the true average from sample averages.

Shaw (1963) showed that punctuation frequency had declined in the 110 years which he included. He chose to make a 3-point study of 1851, 1911, and 1961 publications.[1] If more, and smaller intervals, were chosen, the data could be plotted to show if the rate of decline is constant, accelerating, or decelerating.

Punctuation may, or may not, vary by region. Such variations can be determined by selecting regional publications and making a comparison between them. It is reasonable to hypothesize that the well-known differences in speech in the various regions of the United States may have significant effects upon punctuation. However, a case might be made for the antithesis on the basis of the fact that many people who write for newspapers receive similar educations in the use of punctuation.

The selection of newspapers offers some advantages and some disadvantages. It would be difficult to determine what constituted an adequate sample of printed matter in the United States. To be representative of the whole, the sample would need to include a wide range of printed matter that is published and would require a sample size beyond

practicality. Because this study must be limited because of time constraints, it should be limited to its advantage.

The selection of newspapers has the advantage that they pervade society. There were 615 English language Sunday newspapers in the United States which circulated 50,636,808 copies in 1974. This is one Sunday newspaper for every 4.2 Americans (Delury 1976).[2] Greenberg and Dervin (1970) found that 86 per cent of the families in their general population sample have newspapers delivered to their homes on a daily basis (Lansing, Mich., $n = 206$, $p < 0.01$). They found that the respondents depended upon newspapers for 34 per cent of their world news, and 41 per cent of their local news ($n = 206$, $p < 0.001$).[3] Sargent and Stempel (1968) found that 75 per cent of their general population sample spent 20 minutes or more each day reading newspapers (Athens, O., $n = 114$).[4] Many people do not read a significant number of books, and magazines have selective readership. It is quite certain that--although exposed to other printed matter--most Americans are exposed to the punctuation used in newspapers to a substantial degree.

ENDNOTES FOR CHAPTER III.

[1]Harry Shaw, Punctuate It Right!, New York: Barnes & Noble, Inc., 1963, pp. 11-13.

[2]George E. Delury, ed., The World Almanac and Book of Facts, New York: Newspaper Enterprise Association, Inc., 1976, pp. 201, 565.

[3]B. Greenberg and B. Dervin, "Mass Communication Among Urban Poor," Public Opinion Quarterly, vol. 34 (1970), p. 228.

[4]S. W. Sargent and G. H. Stemple, "Poverty, Alienation, and Mass Media Use," Journalism Quarterly, vol. 45 (1968), p. 325.

CHAPTER IV. METHODOLOGY

The first news page of a Sunday edition of leading newspapers in five regionally distributed American metropolitan areas has been surveyed over four decades. The four years chosen are 1946, 1956, 1966, and 1976. The year 1976 was chosen so that the study would be as current as possible. Ten-year intervals were chosen rather than shorter intervals because Shaw's data indicate a rather slow change in punctuation frequency. It was believed that 10-year intervals would provide adequate trend data if such trends exist. Table III identifies the five areas, their 1970 census populations, the selected major newspaper, and the newspapers' Sunday circulation. The metropolitan areas are not the most populous, but were selected so that the regional difference (if any) may be observed. This selection allows for analysis by significantly different populations in the metropolitan areas. Atlanta and Denver have significantly smaller populations than New York, Los Angeles, and Chicago. Upon approval of the thesis proposal on which this thesis is based, samples of the five newspapers were obtained for the next following Sunday and for the four decade intervals preceding that day. These dates were February 29, 1976, February 27, 1966, February 26, 1956, and March 3, 1946.

TABLE III

MAJOR NEWSPAPERS IN FIVE DISTRIBUTED AMERICAN CITIES[1]

RANK	STANDARD METROPOLITAN AREAS	1970 CENSUS POPULATION	A LEADING NEWSPAPER	SUNDAY CIRCULATION
1	New York	11,528,649	The New York Times	1,507,636
2	Los Angeles	7,032,075	The Los Angeles Times	1,196,143
3	Chicago	6,978,947	The Chicago Tribune	1,005,120
20	Atlanta	1,390,164	The Atlanta Journal & Constitution	580,492
27	Denver	1,227,529	The Denver Post	362,262

IV-1 Preliminary Sample

IV-1.1 Front Page

The first 300 words appearing in the right-hand column of the front page of the 1976 issue of each of the five newspapers were enumerated (n = 1500) (q.v. IV-3, Sample Enumeration Method). This particular portion of the newspapers was chosen because it has large readership. Greenberg and Dervin (1970) found that 60 percent of their general population sample were, ". . . more likely to read the front page . . . ." than the sports, comics, or other sections of newspapers.[2]

The thesis proposal upon which this thesis is based stated that a preliminary sample of 100 words would be taken from each of the five newspapers (n = 500) and that a sample size calculation would be made using the method described by Rickmers (1967)[2]. When this proposed sample was taken, it became evident that this size was inadequate because of the infrequent occurrence of marks of punctuation and the sample was increased to 300 words. The mean frequency of occurrence per kiloword (1000 words), the standard deviation, and a +/- percent confidence intervals based on a statistical probability of 95% (α = .05) were calculated (q.v. Appendix F). These data are presented in Appendix D, Tables D-1 through D-5. In these five tables, the most frequently occurring use of a mark of punctuation, the period (Los Angeles, all uses, $\Sigma X = 20$), results in a mean frequency of occurrence of 66.67 words per kiloword with a standard deviation of 249.9, and +/- percent confidence interval of 42.41. Obviously significant results can be obtained from such a sample, but it is more obvious that significant results cannot be obtained from zero data. Appendix D Tables D-1 through D-5 do not show a single colon, ellipsis, exclamation point, or question mark, and only one semicolon was disclosed. On the basis of this analysis of these preliminary sample data, it was decided to use the methods described in the Curriculum Formulae, Appendix F.

IV-1.2 Second Page

Following the taking of the 300-word samples from the five 1976 sample pages, the entire first pages of each of the 1976 sample pages were enumerated (q.v. Section IV-3, Sample Enumeration Method). The

results of this enumeration are shown in Appendix D, Tables D-6 through D-10 and is discussed in this chapter (IV) under the heading Main Sample (IV-2). The efficacy of continuing on to the second page was considered and samples were taken from the second pages of the 1976 samples. Table IV shows comparative data for the five newspapers between the first and second pages. Hypotheses tests of the differences between mean punctuation frequency were conducted at a probability level of 95% ($\alpha = 0.05$) (q.v. Curriculum Formulae, Appendix F). The results were inconclusive. The Denver sample showed significance, while the other newspapers did not.

In consideration of (1) the results of the above analyses, (2) the acceptability of the level of confidence interval upon the mean in the main sample using page one data only, and (3) time constraints, it was decided to limit the main sample to the entire first page of the twenty newspaper pages represented in this 4 by 5 cell experiment. (This decision was borne out by the analyses of the data collected by these means. Reasonable confidence intervals were found for most marks, but zero occurrences within a given cell occurred for some marks and prevented analysis. The least frequently occurring mark of punctuation in the study was the ellipsis. There were a total of only 5 ellipses in the sample of 52,184 words. These data result in a confidence interval of $\pm$ 87.65% which may be useful. To obtain this level of tolerance within each of the twenty cells of the experiment, it would be necessary to have each cell with a sample size equal to that of the entire experiment. Such a sample would likely comprise the entire first section of each newspaper and would have resulted in a total sample size of approximately one million words. Approximately one day was used to enumerate

TABLE IV

COMPARISON OF PUNCTUATION FREQUENCY ON PAGES 1 AND 2 OF 1976 NEWSPAPER SAMPLES

NEWSPAPER	PAGE 1			PAGE 2			HYPOTHESIS TEST STATS.		
	n_1	$\bar{X}$/kw	s/kw	n_1	$\bar{X}$/kw	s/kw	S_p	t	$t_{\alpha=0.05}$
The New York Times	2134	210.3	464.1	796	202.1	496.8	0.4730	0.4174	1.960
The Los Angeles Times	2200	192.6	431.9	922	173.3	406.4	0.4244	1.159	1.960
The Chicago Tribune	2106	218.3	466.1	865	236.7	500.3	0.4763	-0.9566	1.960
The Atlanta J'rnl & Const.	1648	193.5	442.9	1352	203.3	459.2	0.4502	-0.5932	1.960
The Denver Post	2670	195.1	453.6	341	258.1	482.9	0.4569	-2.396*	1.960

*statistically significant at α =0.05.

each of the twenty cells. At this rate it would take an estimated 400 working days to have taken such a sample. If the data reduction and analyses time were added to this, it would have required two years to complete such a work while working on it full-time.)

IV-2 Main Sample

A main sample comprising the entire first page was taken in each of the twenty cells of this study. It continued from where the preliminary sample of the first page ended (n = 300) and continued until all words (except as explained in Appendix C) on the first page of the twenty newspaper pages had been enumerated and their associated punctuation identified.

The main sample consists of 52,184 words which have associated with them 10,360 marks of punctuation in 2,203 complete sentences. Table V shows the number of words, marks of punctuation, number of sentences, and the number of words per sentence within each of the twenty cells of the study.

Within each of the twenty cells of the study, the occurrence of each of the twelve marks of punctuation and 23 use subcategories was enumerated. The results of these enumerations are shown in Appendix D in Tables D-6 through D-10, and D-16 through D-30.

Each of these twenty tables shows the mean occurrence, the standard deviation, and the +/- percent confidence intervals for the mean based on a probability level of 95% (α = .05). The method of calculation is explained in Appendix F, Curriculum Formulae. The three statistics listed for each mark of punctuation shown in the tables in Appendix D should be interpreted as follows using the Atlanta Journal & Constitution

for February 29, 1976 (Appendix D, Table D-9) as an example. The five lines pertaining to hyphens are interpreted as follows:

> Ninety-five percent (α = 0.05) of all hyphens within another sample of which the Atlanta Journal & Constitution for February 29, 1976 is representative will be within the range of plus and minus 22.9 percent of the sample's mean of 42.5 hyphens per 1000 words (42.5 hyphens/kiloword $\pm$22.9%).
>
> No hyphens were used for compounding words and the means standard deviations and confidence intervals are zero.
>
> Ninety-five percent of all hyphens used to form unit modifiers within another sample of which this one is representative will occur at a frequency of 7.2/kiloword $\pm$56.4%.
>
> Ninety-five percent of all hyphens used for dividing words at ends of lines within another sample of which this one is representative will occur at a frequency of 34.6/ kiloword $\pm$25.5%.

TABLE V

MAJOR GRAPHOLOGICAL CHARACTERISTICS OF MAIN SAMPLE SELEECTIONS

NEWSPAPER	SAMPLE YEAR	TOTAL WORDS	TOTAL PUNCT. MARKS	TOTAL NO. OF SENT.	WORDS PER SENT.
The	1976	2135	449	68	31.40
New	1966	2377	502	94	25.29
York	1956	2810	595	126	22.30
Times	1946	3949	741	149	26.50
The	1976	2201	424	78	28.22
Los	1966	2773	428	95	23.93
Angeles	1956	2459	520	119	20.66
Times	1946	3210	655	120	26.75
The	1976	2107	460	103	20.46
Chicago	1966	2011	405	89	22.60
Tribune	1956	2344	489	114	20.56
	1946	3614	679	145	24.92
The	1976	1649	319	88	18.74
Atlanta	1966	2378	427	115	20.68
Journal &	1956	3132	669	163	19.21
Constitution	1946	3131	612	128	24.46
The	1976	2671	521	108	24.73
Denver	1966	1681	342	65	25.86
Post	1956	2445	373	98	24.95
	1946	3607	750	138	26.14
	TOTALS	52184	10360	2203	$\bar{X}$ = 23.69

The twenty front pages of these newspapers contained 216 stories of which 38 percent were by-line stories, 38 percent were news service stories and 24 percent were either anonymous or identified as such sources as 'Post Staff," 'Times editors," etc. Several pseudonyms were disclosed and these were treated as by-line stories. The content of these stories is discussed in the Epilog to this thesis.

TABLE VI

MAJOR CONTENT OF MAIN SAMPLE SELECTIONS

NEWSPAPER	SAMPLE YEAR	BY-LINE STORIES	NEWS SERVICE	OTHER STORIES	TOTAL STORIES
The	1976	9	-	1	10
New	1966	8	1	2	11
York	1956	7	2	3	12
Times	1946	9	1	3	13
The	1976	7	1	-	8
Los	1966	7	-	-	7
Angeles	1956	2	11	4	17
Times	1946	1	18	2	21
The	1976	4	-	3	7
Chicago	1966	4	1	2	7
Tribune	1956	2	5	4	11
	1946	2	5	4	11
The	1976	4	1	1	6
Atlanta	1966	2	5	3	10
Journal &	1956	3	7	4	14
Constitution	1946	3	6	3	12
The	1976	4	1	2	7
Denver	1966	1	3	4	8
Post	1956	1	5	3	9
	1946	2	9	4	15
	TOTALS	82	82	52	216

IV-3 Sample Enumeration Method

The first pages of the main news section of the four different issues of the five different newspapers were obtained from library and commercial sources. The five issues of February 29, 1976 were actual prints while the others were various kinds and sizes of photographic copies. A form was developed entitled Raw Data Table like the one shown in Figure 1. Data from the newspapers were initially collected on such a form. More than 400 such forms were used and have been bound into four volumes with the newspaper samples for future reference. Summary Data Tables like the one shown in Figure 2 were used to summarize the data for each cell of the experiment including the preliminary samples for a total of 30 tables. The data from these tables were entered as data for the computer program shown in Appendix E.

Each newspaper page was examined under a magnifying glass beginning with the first words of the first story beginning nearest to the upper right-hand column of each sample newspaper. Each word in each article was counted. The first word in each story and each even hundredth word was numbered on the print. Each word that had one or more of the twelve marks of punctuation associated with it was numbered. (There were some exceptions to this method because some of these samples were on reversed image prints and had insufficient 'tooth' to accept writing.)

RAW DATA TABLE for words through . Source:

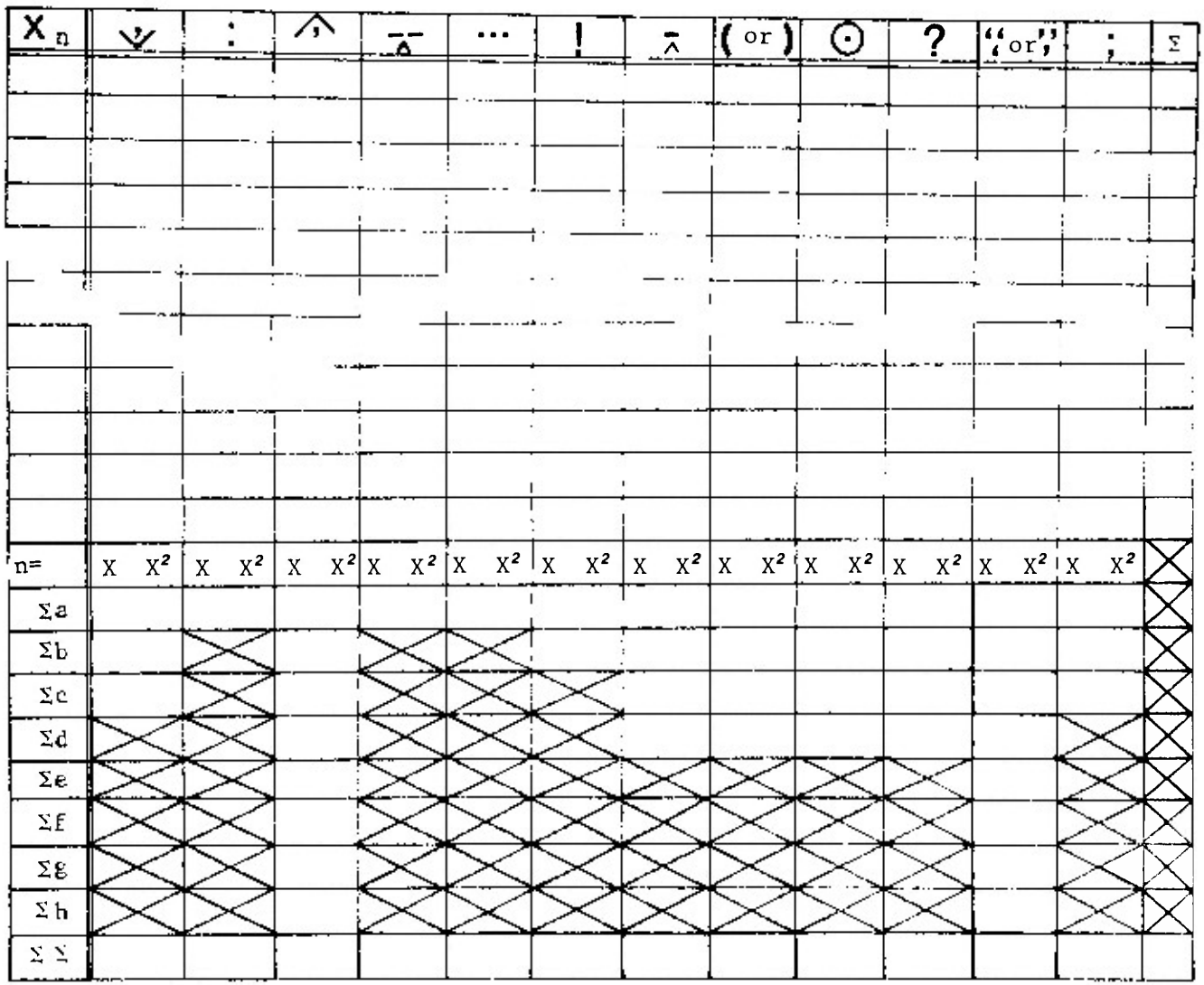


Figure 1. Example of Raw Data Table

Marks of punctuation were counted as associated with the word to which they were placed without spacing. Where marks were placed without spacing with two or more words (e.g., with unit modifiers, dashes, etc.) the first word was used. Compound words were treated as single words so hyphens used with them were treated as intraword marks. Arabic numbers were treated as single words, and decimal points and commas used to divide thousands were treated as intraword punctuation, e.g., 52,183. Where it was not clear if a word was a unit modifier or a compound word (often a question of historical development and not very amenable to objective considerations), the U.S. Government Printing Office Style Guide was used as the standard.[1]

The number of each word which had one or more marks of punctuation associated with it was recorded in the X_n column on the Raw Data Table forms. A use letter was then placed in the column for the particular mark of punctuation. The use letters and the criteria for their use is explained in Appendix C. At the bottom of each Raw Data Table the sum of each use category of occurrence was written (ΣX). The sum of the squares of each occurrence (ΣX^2) was also written at the bottom of each Raw Data Table. When this occurred, the square was different. Often--particularly with commas--the mark performed more than one function. In such cases the sums for the uses exceeds the total sum.

When the enumeration of words and punctuation for each cell was completed, a Summary Data Table was used to sum the sums of X's and

and X's squared and to collect other pertinent data. Figure 2 includes the total number of words in the sample (n), the total number of marks (ΣX), and the total number of individual marks associated with each word squared (ΣX^2). In addition, the number of byline stories, news service stories and the total number of stories was recorded to provide descriptive data about the twenty cells.

These data provided the information necessary to compare the effects of time and location upon punctuation frequency on a cell basis.

SUMMARY DATA TABLE

for: _____ _____

Total number of words in sample (n) _____

Number of byline stories _____

Number of news service stories _____

Total number of stories _____

Total number of punctuation marks (ΣX) quared (ΣX^2) _____

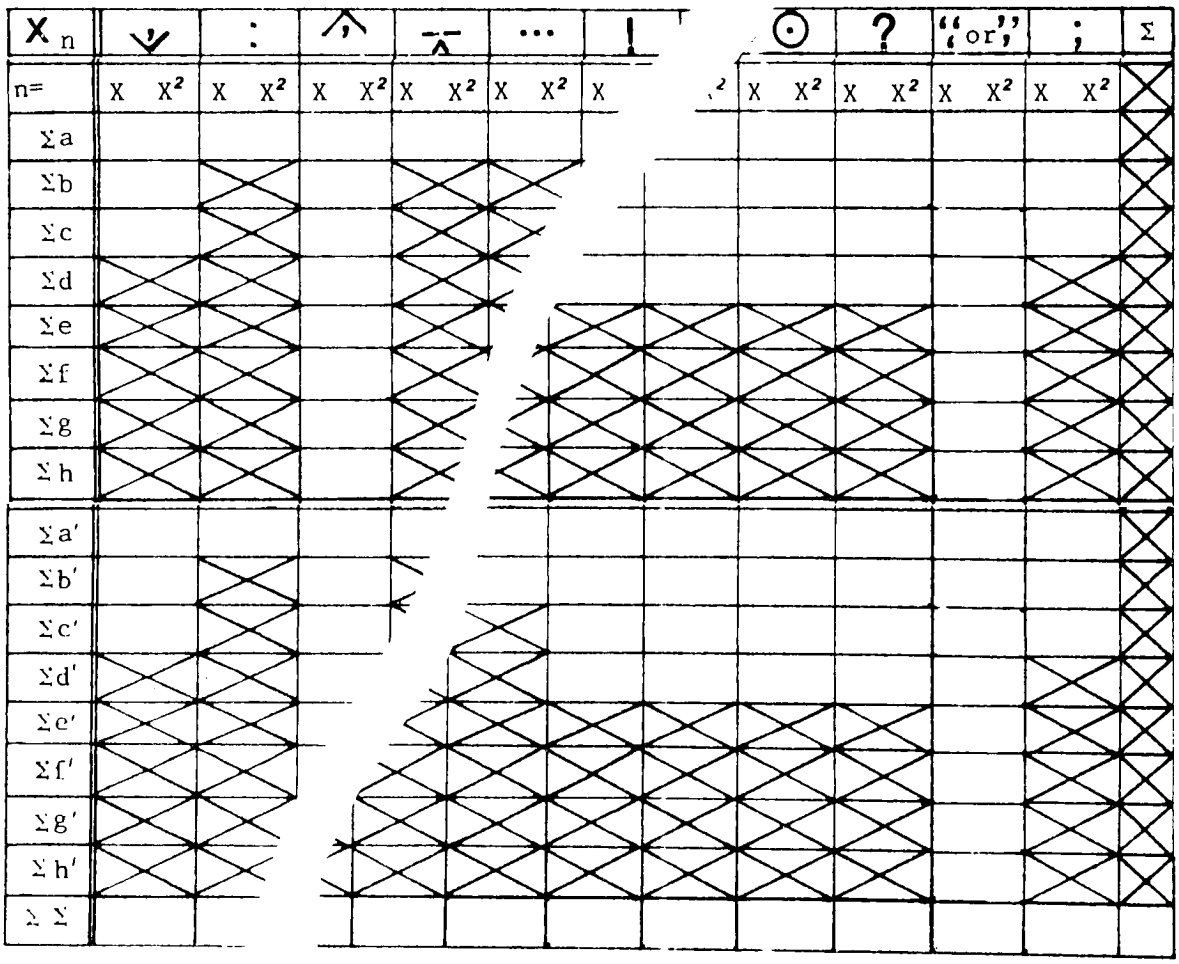


Figure 2. Example of Summary Data Table

ENDNOTES FOR CHAPTER IV

[1]Luman H. Long, ed., The World Almanac, New York: Newspaper Enterprise Association, Inc., 1971, p. 406 and 1973, p. 142.

[2]B. Greenberg and B. Dervin, "Mass Communication Among Urban Poor," Public Opinion Quarterly, vol. 34 (1940), p. 228.

CHAPTER V. HYPOTHESES

1. It is hypothesized that punctuation appears regularly in the main news section of American daily newspapers and that examination of such newspapers will reveal the periodicity of punctuation.
2. It is hypothesized that statistically significant differences in punctuation frequency exist between (a) each of the five selected newspapers, and (b) between the three large-city and two small-city newspapers.
3. It is hypothesized that statistically significant differences in punctuation frequency exist between the decades 1976, 1966, 1956, and 1946 within the selected American daily newspapers.

CHAPTER VI RESULTS

VI-1. Periodicity

Tables D-6 through D-10 and D-16 through D-30 show the frequency of the twelve marks of punctuation and the twenty-three subcategories of use in the twenty cells of this study. Table VII gives the location of the intracell punctuation frequency tables which are to be found in Appendix D. These tables show periodicity ranging from zero (no occurrences of punctuation were found in more than 100 categories out of the 700 reported) to 111 per kiloword as in the case of all commas for the Chicago Tribune sample for 1966. These tables also show the standard deviation and 95% confidence interval, for each of the 700 means. The standard deviations have been coded in the same manner as the means,

TABLE VII

LOCATION OF INTRACELL PUNCTUATION FREQUENCY TABLES

NEWSPAPER	TABLE AND PAGE NUMBERS FOR YEARS			
	1976	1966	1956	1946
N. Y. Times	D-6,101	D-16,111	D-21,116	D-26,121
L. A. Times	D-7,102	D-17,112	D-22,117	D-27,122
Chicago Trib.	D-8,103	D-18,113	D-23,118	D-28,123
Atlanta J & C	D-9,104	D-19,114	D-24,119	D-29,124
Denver Post	D-10,105	D-20,115	D-25,120	D-30,125

i.e., using a kiloword base. The 95% confidence intervals as a percentage of the mean vary from ± 12.41% in the case of the Chicago Tribune for 1966 (q.v. Table D-18) to ± 196% in a number of cases.

VI-2 Decade and Newspaper Differences

Table VIII shows the frequency of all marks of punctuation per kiloword within the twenty cells of this study, and 95% confidence intervals for the five regionally selected newspapers for the three decades between 1946 and 1976. Table VIII shows the means and confidence intervals for all newspapers by decades, for all decades by newspapers, and for all decades by relative city size.

Table IX shows the words per sentence within the twenty cells of this study, and 95% confidence intervals for the five regionally selected newspapers for the three decades between 1946 and 1976. Table IX shows the means and confidence intervals for all newspapers by decades, for all decades by newspapers, and for all decades by relative city size. Words per sentence rather than terminal marks of punctuation per kiloword were used because words per sentence is understood more easily. (Words per sentence is the mathematical reciprocal of terminal marks per word.)

Table X shows same statistics as Table VIII for intrasentence marks of punctuation.

TABLE VIII

MEANS AND 95% CONFIDENCE INTERVALS FOR ALL MARKS OF PUNCTUATION PER KILOWORD FOR THE FIVE SELECTED NEWSPAPERS FOR THE THREE DECADES BETWEEN 1949 AND 1976

SAMPLE YEAR	THE NEW YORK TIMES	THE LOS ANGELES TIMES	THE CHICAGO TRIBUNE	THE ATLANTA J & C	THE DENVER POST	ALL FIVE NEWSPAPERS
1976	210 ± 9.4%	193 ± 9.4%	218 ± 9.1%	193 ± 11%	195 ± 8.8%	202 ± 4.2%
1966	211 ± 8.8%	188 ± 10%	201 ± 9.9%	180 ± 9.7%	180 ± 9.7%	196 ± 4.4%
1956	212 ± 8.0%	211 ± 8.6%	209 ± 9.0%	214 ± 7.7%	153 ± 10%	201 ± 3.8%
1946	188 ± 7.4%	204 ± 7.6%	188 ± 7.6%	195 ± 8.0%	208 ± 7.0%	196 ± 3.4%
ALL DECADES	202 ± 4.2%	200 ± 4.4%	202 ± 4.4%	197 ± 4.4%	191 ± 4.4%	199 ± 1.9%
	200 ± 2.5%			194 ± 3.1%		

TABLE IX

WORDS PER SENTENCE AND 95% CONFIDENCE INTERVALS FOR THE FIVE SELECTED NEWSPAPERS FOR THE THREE DECADES BETWEEN 1949, 1956, 1966, AND 1977

SAMPLE YEAR	THE NEW YORK TIMES	THE LOS ANGELES TIMES	THE CHICAGO TRIBUNE	THE ATLANTA J & C	THE DENVER POST	ALL FIVE NEWSPAPERS
1976	31 ± 23%	28 ± 22%	20 ± 19%	19 ± 20%	25 ± 18%	24 ± 9.1%
1966	25 ± 20%	24 ± 20%	23 ± 20%	21 ± 18%	26 ± 24%	23 ± 9.0%
1956	22 ± 17%	21 ± 18%	21 ± 18%	19 ± 15%	25 ± 19%	21 ± 7.7%
1946	27 ± 16%	27 ± 16%	25 ± 16%	24 ± 17%	26 ± 16%	26 ± 7.4%
ALL DECADES	26 ± 9.2%	25 ± 9.5%	22 ± 9.0%	21 ± 8.6%	25 ± 9.5%	24 ± 4.1%
	24 ± 5.3%			23 ± 6.4%		

TABLE X

MEANS AND 95% CONFIDENCE INTERVALS FOR INTRASENTENCE MARKS OF PUNCTUATION PER KILOWORD FOR THE FIVE SELECTED NEWSPAPERS FOR THE THREE DECADES BETWEEN 1946 AND 1976

SAMPLE YEAR	THE NEW YORK TIMES	THE LOS ANGELES TIMES	THE CHICAGO TRIBUNE	THE ATLANTA J & C	THE DENVER POST	ALL FIVE NEWSPAPERS
1976	$178 \pm 11\%$	$157 \pm 11\%$	$169 \pm 11\%$	$140 \pm 14\%$	$155 \pm 10\%$	$161 \pm 5.0\%$
1966	$172 \pm 10\%$	$147 \pm 12\%$	$157 \pm 12\%$	$131 \pm 12\%$	$165 \pm 12\%$	$154 \pm 5.2\%$
1956	$167 \pm 9.5\%$	$163 \pm 10\%$	$160 \pm 11\%$	$162 \pm 9.4\%$	$112 \pm 13\%$	$154 \pm 4.6\%$
1946	$150 \pm 8.6\%$	$167 \pm 8.8\%$	$148 \pm 9.0\%$	$155 \pm 9.5\%$	$170 \pm 8.1\%$	$157 \pm 3.9\%$
ALL DECADES	$164 \pm 4.8\%$	$159 \pm 5.1\%$	$157 \pm 5.2\%$	$149 \pm 5.4\%$	$152 \pm 5.2\%$	$156 \pm 2.3\%$
	$160 \pm 2.9\%$			$150 \pm 3.7\%$		

CHAPTER VII. ANALYSIS OF DATA

VII-1 Individual Cells

As stated previously, the sum of the occurrences (ΣX) and the sum of the occurrences squared (ΣX^2) were calculated within each of the twenty cells of this study for the twelve marks of punctuation and the 23 subcategories of use and used to generate the tables in Appendix D. These seventy sums (35 ΣX's and 35 ΣX^2's) together with the number of words in each cell (n) were used to calculate 35 means (punctuation frequencies), 35 standard deviations, and 35 confidence intervals as a $\pm$ percentage of the mean based on a probability level of 95% (α = 0.05). Appendix F explains the means by which the calculations were made. Section IV-2 of Chapter IV explains the interpretation of these statistics. These 105 statistics are presented for each of the twenty cells of this study in Appendix D (q.v. Table VII).

The data presented in Appendix D is *prima facie* evidence confirming the hypothesis of periodicity with the exception of those instances where no occurrences were found and those were very few occurrences were found. Using the confidence interval technique provides a quality value for the reliability of the frequency statistics.

VII-2 Across Cells

As previously stated, the sum of the occurrences (ΣX) and the sum of the occurrences squared (ΣX^2) were calculated for the punctuation

frequency for each of the twenty cells in this study. Also, the sums of the occurrences (ΣX) and the sums of the occurrences squared ($\Sigma X^{2)}$ were calculated for terminal marks of punctuation. From these data, the number of intrasentence marks of punctuation were calculated by subtracting the number of terminal marks in each cell from the total number of marks for both the sum of the occurrences (ΣX) and the sum of the occurrences squared (ΣX^2). This produced a data matrix of 140 discrete numbers. These data are shown in Appendix E, Tables E-1, E-2, and E-3. From these data, means and standard deviations were calculated for each cell of this study, and also clustered groups by decade and newspaper. These statistical data are shown in Tables VIII, IX and X in Chapter VI.

To test the statistical significance of the data presented in Table VIII, six analyses of variance were performed. These analyses were designed to test hypotheses 2 and 3 (q. v. Chapter V). Table XI, XII and XIII show the statistical data by both decade and newspaper for all marks of punctuation, terminal marks of punctuation, and intrasentence marks of punctuation respectively. Pearson's "F" distribution test of significance was used as detailed in Appendix F.

TABLE XI

EFFECTS OF DECADE AND NEWSPAPER ON TOTAL MARKS OF PUNCTUATION

EFFECT		SUM OF SQUARES	DEGREES OF FREEDOM (ν)	MEAN SQUARES	F STATISTIC	F_α
TIME	Between Groups	0.3225	3	0.1075	0.5302	2.6049 (α = 0.05)
	Residual (Error	10580.9245	53,180	0.2028	-	-
	Total	10581.2470	-	-	-	-
REGION	Between Groups	0.9713	4	0.2428	1.1976	2.3719 (α = 0.05)
	Residual (Error)	10580.2757	52,179	0.2028	-	-
	Total	10581.2470	-	-	-	-

TABLE XII

EFFECTS OF TIME AND REGION ON TERMINAL MARKS OF PUNCTUATION

EFFECT		SUM OF SQUARES	DEGREES OF FREEDOM (ν)	MEAN SQUARES	F STATISTIC	F_α
TIME	Between Groups	0.5139	3	0.1713	4.2373**	3.7816 (α = 0.01)
	Residual (Error)	2109.4842	52,180	0.0404	-	-
	Total	2109.9981	-	-	-	-
REGION	Between Groups	0.6577	4	0.1644	4.0672**	3.3192 (α = 0.01)
	Residual (Error)	2109.3405	52,179	0.0404	-	-
	Total	2109.9981		-	-	-

**Statistically significant at α = 0.01

TABLE XIII

EFFECTS OF TIME AND REGION ON INTRASENTENCE MARKS OF PUNCTUATION

EFFECT		SUM OF SQUARES	DEGREES OF FREEDOM (ν)	MEAN SQUARES	F STATISTIC	F_α
TIME	Between Groups	0.2199	3	0.0732	0.4180	2.6049 (α = 0.05)
	Residual (Error)	9141.1679	52,180	0.1752	-	-
	Total	9141.3876	-	-	-	-
REGION	Between Groups	1.5886	4	0.3971	2.2673	2.3719 (α = 0.05)
	Residual (Error)	9139.7991	52,179	0.1752	-	-
	Total	9141.3876		-	-	-

CHAPTER VIII. DISCUSSION OF RESULTS

VIII-1. Periodicity of Punctuation

Chapter VI and Appendix D show the frequency of the twelve marks of punctuation and the 23 subcategories of uses. The confidence interval given for each frequency provides a quality value for these means. At first it might seem that punctuation frequencies with 95% confidence intervals of ± 40%, ± 50%, ± 100% or more might not be very useful. Actually, they can be very useful in the appropriate circumstance when applied correctly.

The writer who wishes to compare his work with these data would first check his material for the number of words per sentence shown in Table IX. This will tell him if his sample is representative of the population from which the samples were taken for this study, i.e., the various newspapers and various decades--presumably he would be interested in the current decade. If he finds that his work is not in consonance with these ranges, further comparisons would be invalid because different numbers of words per sentence will likely result in a different distribution of punctuation frequencies.

The writer who finds his work within the words-per-sentence ranges would proceed to compare his work with the frequency rates for all intrasentence marks of punctuation as shown in Table X. The confidence intervals for these frequency rates are rather narrow. (The mean frequency for all 1976 samples is 161 marks per kiloword with a 95% confi-

dence interval of $\pm$ 5%.) If he finds that his punctuation frequency is variant, he may investigate by mark, and then by use of that mark. The 95% confidence interval become wider as one investigates in more detail, but it should be possible to determine a writer's differing use of punctuation by these means.

The 95% confidence intervals may seem rather wide for other applications, but it should be remembered that the literature search yielded no other body of comparable data to that presented in this thesis.

VIII-2 Decade and Newspaper Differences

Tables VIII, IX, and X show the decade and newspaper difference of punctuation frequency for all marks, terminal marks, and intrasentence marks, respectively. Tables XI, XII, and XIII show statistical analyses of variance of the punctuation frequencies shown in Tables VIII, IX, and X. Only one difference is statistically significant, viz., the effect of decade and newspaper on the number of terminal marks of punctuation which is significant at a probability level greater than 99% ($\alpha = 0.01$). (All other differences are not statistically significant at a probability level of 95%.)

Figure 3 shows a graphical representation of the trend exhibited by sentence length. It would appear that sentence length declined in the post-World War II era and then began to rise at the time of the War in Viet Nam. It is apparent that the trend is not a simple increase or decrease and, with the level of significance obtained for these data, it is unlikely that this is a chance occurence.

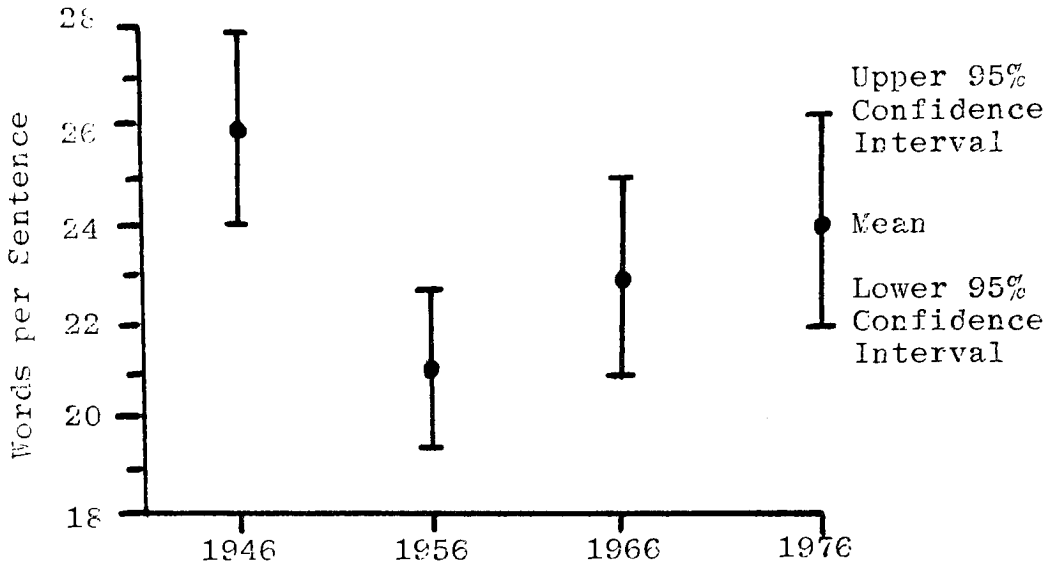


Figure 3. Trend in Sentence Length as a Function of Decade

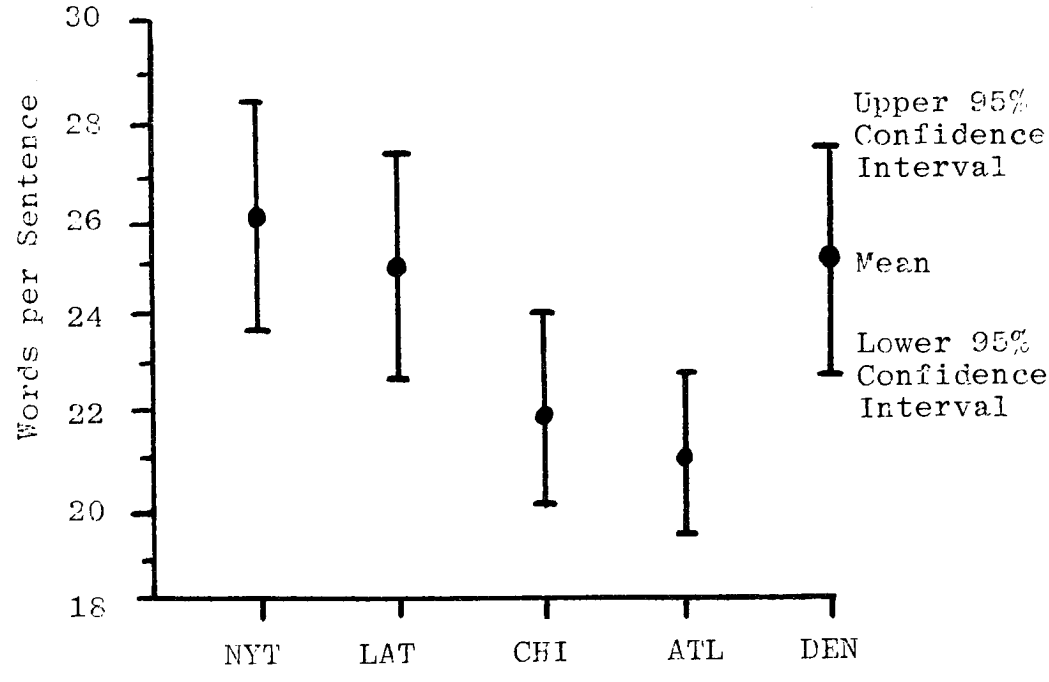


Figure 4. Trend in Sentence Length as a Function of Newspaper

The decline in use of punctuation reported by Shaw (q.v. Chapter II) has not been sustained by this study. Shaw did not consider the length of sentences and he did not report sufficient information to replicate his study, so direct comparison of this study and Shaw's study is not possible without further extensive research.

CHAPTER IX. SUMMARY AND CONCLUSION

This thesis has reported on the study of twelve marks of punctuation and twenty-three subcategories of use across the three decades between 1946 and 1976, and for five regionally selected newspapers. The character of the samples is shown in Table V and Table VI. The twelve marks and the twenty-three subcategories are explained in detail in Appendix C.

The frequency of punctuation by mark and use of mark for each of the twenty cells of this study are given in a series of computer-generator tables in Appendix D. Punctuation frequencies vary from no occurrences within the sample to 111 commas per kiloword in one instance. Ninety-five percent confidence intervals vary from $\pm$ 12.41% to $\pm$ 196% as a percentage of the mean. Six hundred of the seven hundred punctuation frequencies occur with rates which are useful.

The overall frequency of all, terminal, and intra-sentence punctuation was also studied. It was found that the frequency of all punctuation was 199 marks/kiloword $\pm$1.9%, the mean length of sentences was 24 words per sentence $\pm$4.1% (the reciprocal of terminal marks per word), and the frequency of intrasentence punctuation was 156 marks/kiloword $\pm$2.3%

Analyses of variance were conducted to find if regional and/or historic trends could be found. No statistically significant differences were found for all punctuation or for intrasentence punctuation, but sentence length was significantly different both as a factor of time and newspaper.

The trend in sentence length was neither a simple rising nor falling trend. Figures 3 and 4 display the data in graph form. With one exception (The Chicago Tribune), sentences in 1956 were shorter than other years. No hypothesis is offered to explain this result. More data would be required to either describe the trend or postulate an explanation for the trend.

CHAPTER X. RECOMMENDATIONS

It is believed that the data presented is conclusive of the issue of trends in punctuation frequency over the three decades considered, i.e., that the frequency is not changing. However, the issue of change in sentence length is not conclusive. Certainly the data presented here could be analyzed further to search for causes of the trend shown, and most certainly more data would provide further information on the trend.

Beginning with the four time points in this study, one could make a 30-point study by sampling each year in the end of February. Samples would need be taken only of the total number of words and the total number of sentences and these could be estimated with reliability with somewhat smaller samples than those used for this study.

* * * *

A plan for the use of the data as a normalized standards for writers could be developed. This is alluded to in Chapter VIII. A computer program could be developed which would analyze and compare a person's punctuation behavior with the standard frequencies included herein.

* * * *

During the collection of data for this study, the possibility of collecting type line length and comparing it to the frequency of word-break hyphens was considered. The variety of typography did not make that a reasonable analysis to consider including here. However, it is

easy to see how such a study might be designed and how printers might make use of nomographs which predict the number of word-break hyphens that would result from various combinations of type measures, type faces, and type points.

CHAPTER XI. EPILOG

In this thesis, much has been written about the frequency of punctuation in the samples and the samples have been described in detail. Samples were taken over three decades of history and nothing has been said so far about the information that these newspaper pages conveyed. In this chapter, there will be a few words besides the words (the meaning of epilog) about punctuation frequency. The following comments are presented anti-chronologically i.e., in the same way the study was developed.

February 29, 1976

At this writing, it seems strange to realize that just over a year ago the newspapers considered President Ford as the front-runner for the presidency. Ronald Reagan and President Ford were campaigning in Florida and in Massachusetts, "The Democratic candidates in [this] crowded, confused field [were] running in Tuesday's Presidential primary ...." the New York Times reported. The Los Angeles Times had the headline, "Wallace Leads Carter in S.C," while The Atlanta Journal & Constitution had its lead story headlined, "Wallace Edges by Carter." The front page of the Denver Post did not even mention the presidential election primaries!

February 27, 1966

The New York Times lead off with three stories about President Johnson. The Chicago Tribune quoted the President as saying, "It [the

Viet Nam War] is going to be difficult and it is going to require sacrifices. We want everyone to know that." And, ". . . it's not going to be easy and it's not going to be short."

The newspapers were full of space news. The day before, Saturday, February 26, 1966, the Saturn 1B rocket sent an unmanned Apollo spacecraft into suborbital flight. This was an early test for the flight which resulted in the Lunar landing some years later.

February 26, 1956

The time that has elapsed since this date is clear when you see the headlines in the New York Times: "Harriman [Governor of New York] Urges Rise in 'Gas' Tax to pay for Roads," "Democrats Score Dulles' Optimism on Soviet Threat," "Stevenson Gibes at the President as Inept 'Coach,'" and, "Eisenhower Flies Back to Capitol; Decision Waited." The decision referred to in this last headline was whether or not he would run for the presidency again.

March 3, 1946

On this date, several papers headlined a story largely forgotten today. The Chicago Tribune headlined, "Greenwich [Conn.] Votes Against UNO Site." The United Nations--then called the United Nations Organization and not printed with initial capital letters--was looking for a sight for its world headquarters. Greenwich was considered but did not welcome the new organization.

It was not long after the end of World War II as one can tell from these headlines from The Los Angeles Times: "Sgt. Hirt Wins Acquittal by Court-martial," and "Hitler's Wrong Guesses in '42 Helpful to Allies."

* * * *

This brief glimpse into the content of the sample may help give the reader some concept of what a period of thirty years encompasses. It included seven presidents: most of Truman's first term, all of Truman's second term, Eisenhower's eight years, and the terms of Kennedy, Johnson, Nixon, and Ford. It included the entire term of Nelson Rockefeller--New York's longest serving governor. It included all of the Korean and Viet Nam Wars. It included all of America's space program, the invention of the transistor, and the development of the computer. However, in that vast span of time, the frequency of punctuation punction in newspapers has not changed.

BIBLIOGRAPHY

Delury, George E., ed., The World Almanac and Book of Facts, New York: Newspaper Enterprise Association, Inc., 1976, pp. 201, 565.

Greenberg, B., and B. Dervin, "Mass Communication Among Urban Poor," Public Opinion Quarterly, vol. 34 (1970), p. 228.

Johnson, Donald D., Communication Characteristics of NTID Students, Rochester, N.Y.: Rochester Institute of Technology, National Technical Institute for the Deaf, 1976, pp. 25--29.

Lapin, Lawrence L., Statistics for Modern Business Decisions, New York: Harcourt Brace Jovanovich, Inc., 1973.

Long, Luman H., ed., The World Almanac, New York: Newspaper Enterprise Association, Inc., 1971, p. 406, and 1973, p. 143.

Rickmers, Albert D., and Hollis N. Todd, Statistics, An Introduction, New York: McGraw-Hill Book Company, Inc., 1967.

Sargent, S. W. and G. H. Stempel, "Poverty, Alienation, and Mass Media Use," Journalism Quarterly, vol. 45 (1968), p. 325.

Shaw, Harry, Punctuate It Right!, New York Barnes & Noble, Inc., 1963.

Style Manual (abridged), Washington: United States Government Printing Office, 1967.

Zais, Robert S., A Study to Determine Whether a Punctuation Symbol's Characteristics Influence the Effectiveness of a Method of Teaching Punctuation, unpublished Masters thesis, Department of English and Education, Brown University, 1963.

APPENDIX A

A SHORT HISTORY OF PUNCTUATION

by Robert S. Zais*

*Reproduced with permission of the author from the diachronic and synchronic statements of the author's Master's Thesis, 1950 (q.v. Bibliography).

CHAPTER I

INTRODUCTION

A. Diachronic Statement

Although men have been writing for at least six thousand years, punctuation as we know it is only a few hundred years old. The ancient Phonecians and Greeks, who three thousand years ago developed the system of writing that we use today (in which marks stand for sounds rather than words or ideas), used no symbols which correspond to our own punctuation marks; in fact, all writing was done in large letters, and no spaces were used to separate words or structural patterns. "THISISWHATANCIENTGREEKLOOKEDLIKEOFCOURSEITISNTGREEKITSENGLISH"[1]

Priests and preachers of the Middle Ages were the first people to use marks resembling our own punctuation symbols. The purpose of these marks was to indicate the various pauses which the reader was to make in his oral delivery of the manuscript. (Compare today's radio scripts, which make use of a variety of marks-underlines, suspension dots, hyphens, the word "Pause," and various other symbols not used in conventional printing--in order to quide the reader in his delivery).[2] But the "punctuation" in the Mediaeval manuscripts was, as Vallins puts it, "light . . . and . . . haphazard."[3] It wasn't until the invention of printing and its introduction into England in the last quarter of the 15th century that something resembling a system began to emerge. Caxton, the first English printer, used three symbols-" a stroke (/) marking

off word-groups or phrases, a colon marking a distinct syntactical break or pause, and a full stop marking either the end of the sentence or a brief pause."[4] Although there were inconsistencies, it is clear that a general pattern was followed.

By 1530, the use of a period to mark a brief pause was eliminated, and by the end of the century, the semi-colon was introduced and the stroke replaced by the comma. These four marks (period, colon, semi-colon, and comma), formed the basis of 17th century punctuation, for which Bacon's Essays serves as a typical example. It is clear from his arbitrary use of the colon, semi-colon, and comma that punctuation convention of that day was based on breath pauses rather than syntactic structure. Ben Jonson's Grammar (the 1692 revised edition) concurs in this theory of punctuation. "Whereas," he says, "our breath is by nature so short, that we cannot continue without a stay to speake long altogether, it was thought necessarie, as well as for the speaker's ease, as for the plainer deliverance of the things spoken, to invent this meanes, whereby men pausing a pretty while, the whole speech might never the worse be understood."[5] Lowth, another grammarian, also feels that punctuation is the means of writing "pauses or rests" and even assigns mathematical values to each symbol. Thus, "the Period = 2 Colons: the Colon = 2 Semicolons; the Semicolon = 2 Commas."[6] But such inaccurate descriptions of the punctuation system could not help but lead to the chaos in which 18th century writers found themselves. In the words of Vallins (page 72), "Colons, semi-colons, and commas seem to jostle one another unceremoniously, with little regard for their relative values; and the grammarians themselves find some difficulty in distinguishing one from the other."

William Cobbett, recognizing the state into which the punctuation system had fallen, was to relegate the semi-colon to the position of a stop used "when the comma is not quite enough to keep the meaning of the sentence sufficiently distinct."[7] (Note the structural implication of this statement). He then goes on (page 152) to admit the impossibility of stating any precise rules for the use of the four punctuation symbols and calls upon the writer's "taste" as a basis for judgment. "Something must depend [he says] upon the weight which we may wish to give to particular words, or phrases; and something on the seriousness, or the levity, of the subject on which we are writing." Here, Cobbett vaguely indicates the necessity for basing punctuation on something more than measured breathing pauses.

Justin Brenan, a disciple of Cobbett's, is likewise aware of the inadequacy of pause measurement as the basis for punctuation. In Composition and Punctuation familiarly explained (1829), he states that when any mark other than the comma is introduced into a sentence, the "entireness of the sentence is, in some degree, broken."[8] He is contemptuous (page 153) of the "useless controversial stuff [that] has been written upon the "proper" use of the semi-colon and colon." Indeed, he would abolish both and substitute the dash for them. Although Cobbett, his master, had no use for the dash because no one could decide just how long it should be, Brenan defended it as a more "expressive" mark which anyone of good taste would prefer to the colon. He would not, however, indiscriminately replace all colons and semi-colons with dashes, for in citing a sentence in which the writer had used a colon, he says, "Now, I say, that there should be neither colon, nor semi-colon, but a simple

comma . . ." It is not so much the correction that Brenan has made that is interesting or significant, but rather the reason which he gives for making the correction. He says (page 154), "The colon would never have been introduced here, but for that fruitful source of vicious punctuation-the vain attempt to fix portions of time for pausing on each stop."

But although both Cobbett and Brenan felt the inadequacy of "measured pause" punctuation, neither fully recognized punctuation's syntactic significance. Indeed, Brenan's directions on the use of the comma are couched in "pause philosophy." "We may generally put it," he says, "whenever we think there is an ordinary rest before the end of the sentence."[10] As a result of this philosophy and because he preferred what is generally referred to as "close punctuation," we find him offering the following example (page 155):

> Now, as this may, and often has, and is still, considered, by many persons, as a direct infringement, yet, seeing that, under all circumstances, men will, in despite of law, run those risks, here, take into serious account . . . .

Although the "measured pause" philosophy for determining the use of colons, semi-colons, and commas was losing ground, the "pause" philosophy for the use of commas remained strong until the end of the 18th century so that commas often obscurred or even upset the syntactic structure of a sentence. Thus we find instances of relative clauses being set off by commas from main clauses whether they are "restrictive" (defining or modifying) or "non-restrictive" (non-defining or descriptive). A sample offered by Vallins (page 155) is:

> A wretch, who in the deepest distress still aimed at good-humour, was an object my friend was by no means capable of withstanding. --Goldsmith

The relative clause is, of course, a defining modifier of "wretch," and present day usage would have omitted the commas. It is interesting that Lindley Murray attempted to define the distinction between "restrictive" and "non-restrictive" clauses:

> Relative Pronouns are connective words, and generally admit a comma before them; as, 'He preaches sublimely, who lives a sober, righteous, and pious life:' 'There is no charm in the female sex, which can supply the place of virtue.'
>
> But when the two members are closely connected by a relative, restraining the general notion of the antecedent to a particular sense, the comma should be omitted; as 'A man who is of a detracting spirit, will misconstrue the most innocent words that can be put together.'
>
> In this example, the assertion is not of 'a man in general,' but of 'a man who is of a detracting spirit;' and therefore they should not be separated.[11]

Murray is unaware, of course, that the clauses in the first two sentences "restrain the general notion of the antecedent to a particular sense" as much as the clause in the third sentence does. Notice, too, that in the third sentence, where modern usage would never allow a comma separating the subject from the verb, Murray finds its use necessary. Indeed, he makes it a point to comment on this use of the comma. "A simple sentence, when it is a long one, and the nominative case is accompanied with inseparable adjuncts, may admit of a pause immediately before the verb . . . ."[12] Anyone familiar with 18th century literature is aware of the frequency with which this "pause" is indicated by a comma.

If the tendencies of men like Murray and Brennan toward a punctuation system based on syntactic structures rather than breath pauses

was a step in the right direction, their work--because it mingled the old theories with the new--resulted in many illogicalities. The sentence cited above, in which the comma before the relative pronoun is omitted for structural reasons, is a case in point. By modern standards it would seem that the less serious offence would be to admit commas which set off restrictive clauses that to eliminate only the first and thus divide the sentence into two distinct segments--the subject and the predicate. As Vallins so aptly puts it (page 157). "Today we have more respect for the pattern of the sentence; a comma is not allowed to divide parts which are properly and logically indivisible." But in the 18th century, where the comma before the relative pronoun of a restrictive relative clause was omitted on the basis of structure, the comma which separated subject from verb was based on pause.

While contemporary writers and editors might find such practice unsophisticated, even today, many situations involving what are called problems in "open" or "close" punctuation are resolved on the basis of pause. What Vallins warns against (page 157) is "the half-hearted attempt at open punctuation [which] gives a false and unintelligible sentence pattern." His example is the clause ". . . while light conversation, punctuated by squeals of pain is to be expected." In this clause the decision as to whether or not commas will be used to set off the participial "punctuated by squeals of pain" will be determined by the meaning which the writer wishes to convey. But the inept attempt at "open" punctuation illustrated above not only obscures the meaning, but does violence to the structure of the sentence. Thus Vallins observes that while a writer may use stops with a certain freedom, he must recog-

nize "a basic system that has some relationship to syntax" and stay within the limits of that system in formulating his style.

Besides the period, colon, semi-colon, and comma, Cobbett recognizes the question mark and the exclamation mark. But Murray adds the warning that "a note of interrogation should not be employed where it is only said a question has been asked, and where the words are not used as a question."[13] In spite of this warning, the practice of following an indirect question with a question mark was common in the 18th century, and it seems that this convention was more than a little encouraged by the internal punctuation used in writing indirect questions. Vallins illustrates with the sentence "The Cyprians asked me, why I wept." The comma after "me" has the effect of setting off the words following, thus tending to give the grouping the appearance of a direct question. Here we have another example of the mingling of the pause and the structure theories contributing to illogicalities. As Vallins says (page 158), "Modern punctuation, without the comma, has gone far towards abolishing this error."

Although the exclamation mark comes to us accompanied by a variety of interesting little stories (it was once called the "note of admiration"; it originated from "Io," the Latin exclamation for joy), it has no special historical interest. Its use, according to Cobbett, simply indicates an expression of surprise as well as a statement of fact.

Quotation marks, unknown to Ben Jonson, made their appearance later than the marks already discussed. Their use to indicate direct speech was not recognized by early grammarians, or at least, according to Vallins, this use was not stressed. "Two inverted commas," says Murray,

"are generally placed at the beginning of the phrase or passage, which is quoted or transcribed from the speaker or author in his own words; and two commas, in their direct position, are placed at the conclusion."[14] Cobbett's theory closely parallels Murray's except that he describes what we would now call "single quotes." Both men think of the marks as indicating quotations as opposed to, and as distinct from, direct speech.

In 1759, however, we find Samuel Johnson, in his Rasselas, indicating direct speech with quotation marks. Vallins presents several quotations from the work to illustrate certain curiosities in Johnson's style. Most noteworthy is Johnson's tendency to include interpolated expressions such as "said Rasselas" within the quotations in some instances, but (as in modern usage) to fail to do so in others. Another peculiarity is his failure to use some means of distinction (such as single and double quotes) to indicate a quotation within a speech. Although such practices have largely been eliminated by modern punctuating convention, a still unresolved question illustrated in this work is the problem of deciding when stops at the end of quotations are to be placed inside the quotation marks and when they are to be placed outside. Older writers and printers tended to enclose all end punctuation within the quotation without reference to its connection with the quotation or with the sentence. To illustrate, Vallins (page 159) quotes from Rasselas:

> (Imlac is speaking) "I suppose he discovered in me, through the obscurity of the room, some tokens of amazement and doubt, for, after a short pause, he proceeded thus:"
>
> "Not to be easily credited will neither surprise nor offend me; for I am . . ..

Johnson has placed the colon which introduces the subsidiary speech inside the quotation marks; modern usage, which puts stops inside or outside the quotation marks accordingly as they belong to the quotation or to the sentence of which the quotation is a part, would probably have the colon outside the quotation marks. But although some modern writers and publishers (Vallins quotes from The Times Literary Supplement) tend to observe "modern usage" in cases involving colons and semi-colons, they ordinarily include other stops within the quotation marks whether or not they are integral to the quoted material. compare these sentences:

> It is not much good English people objecting to the American mistranslation of Kapellmeister as 'chapel master,' which they now apply to the leader of the orchestra ...
>
> To understand why this should be we must return to Professor Kemp Smith's phrase 'Descartes as Pioneer'; for his importance . . .[15]

It is clear from these two examples (and other similar ones which occur in the same issue of The Times Literary Supplement) that the printers have punctuated on the basis of whim rather than on the basis of some intelligible principle. Moreover, we find that situations often arise in which it is not always easy to determine when a stop is an integral part of the direct speech which is quoted. The complications and difficulties which arise with respect to determining the position of such stops have led Vallins (page 161) to declare, ". . . quotation marks--at any rate for direct speech--should be abolished altogether."

Another punctuation symbol whose "proper use" in many areas is still controversial is the apostrophe. This symbol's history as a mark

used to form the modern English genitive or possessive is recorded in The Oxford English Dictionary. It "originally marked merely the omission of e in writing, as in fox's, James's, and was equally common in the nominative plural, esp. of proper names and foreign words (as folio's = folioes); it was gradually disused in the latter, and extended to all possessives, even where e had not been previously written, as in man's, children's, conscience' sake."[16]

But even as late as the last quarter of the 19th century we see grammarians objecting to the use of the apostrophe to form the possessive of plural nouns (such as birds', boys') "because no vowel has been dropped there."[17]

Absurd as this argument may sound today, it is typical of the attitude which has governed (and to some extent still does govern) thinking with regard to teaching proper punctuation; that tradition and logic determine the laws to which usage must conform. We shall see in the next section that statements by modern linguists have done much to clarify the role of punctuation in written language; and in so doing, they have cleared the way for an experimental approach to determining by which methods punctuation skills are best taught.

B. Synchronic Statement

While the printers of the 16th and 17th centuries did a great deal toward normalizing punctuation and modern editors have made it, as Vallins puts it, "both an art and a science," our lack of understanding of the principles governing the system has led to the difficulties and controversies which "proper punctuation" still involves.

Ephraim Chambers recognized this situation over 200 years ago when he wrote:

> There is much more difficulty in pointing, than people are generally aware of.--In effect, there is scarce any thing in the province of the grammarians so little fixed and ascertained as this. The rules usually laid down are impertinent, dark, and deficient; and the practice, at present perfectly capricious . . .[18]

With the advent of linguistic science, great strides were made in terms of understanding how language works, but for some reason, punctuation was given little consideration, even in studies which undertook to explain the principles underlying writing systems. Such books as Leonard Bloomfield's Language and Stuart Robertson's The Development of Modern English do not mention punctuation in their chapters on written language. Although Sledd mentions punctuation in his book, A Short Introduction to English Grammar, he does so obliquely. In Chapter Eight, "Applied Grammar: Some Notes on English Prose Style," Sledd, introducing an exercise in grammar, states, "None of the following sentences would be ambiguous in speech, but in writing they might be misleading . . . Revise them, or punctuate them [my italics] so that they will be immediately clear."[19] The important role of punctuation is clearly implied, but no detailed linguistic treatment of the subject is provided.

The history of punctuation and the work of some linguists, however, clearly indicates that advances have been made. For instance, although we still recognize the "pause" feature of certain punctuation symbols, we do not allow a syntactic structure to be obscured simply to indicate such pause. It must be admitted, however, that much of the progress we

have made may be the result of improved methods of reading rather than a more profound understanding of punctuation's role. Paul Roberts, among others, points out that 1500 years ago people moved their lips when reading, carefully forming every sound that was represented by the letters; thus punctuation was significantly intonational. Today, since we read with closed lips and try to pick up larger and larger stretches of print with a single movement of the eye, it is important that the printed line be left as open as possible. This trend toward lighter and lighter punctuation is readily observable. Hence, we find such statements as the following, which attempt, in a general way, to explain the role of punctuation in writing:

> Punctuation marks are one means of helping us get our exact meaning on the page. They do more than mark such obvious facts of language as "This is a sentence," "This is a question." They help us separate words (and thoughts) and so present them distinctly to the reader; they help group and keep together related ideas; they set off certain words for emphasis. Their use affects the tempo of writing: Too many marks may slow the reader to the point of exasperation, and too few may make him go over a passage two or three times to get its probable meaning. The writer who wishes his work to appear to the best advantage will give close attention to its punctuation.[20]

> Skillful punctuation--including the omission of useless and obstructive punctuation--makes clear at a glance the relations and sometimes the junctions of word groups, and often indicates their relative weights. A badly chosen mark annoys the reader by asking him to do the writer's work of grouping the material, or distorts the intended emphasis. Omission of a needed mark puzzles the reader momentarily and gives him the trouble of guessing at the writer's intention. Useless marks check the progress of reading and suggest groupings and relations that are not intended.[21]

> Punctuation is in large part a system of conventions the function of which is to assist the written language in indicating those elements of speech which cannot be conveniently set down on paper: chiefly pause, pitch, and stress. It is relevant here that the words period, colon, and comma all signified a sentence, a portion of a sentence, or a pause, long before they came to be applied to the various points of punctuation. Such points as the question mark and the period are obvious substitutes or compensations for pitch modulation; the exclamation point suggests the element of stress or volume; frequently the comma and semicolon correspond to pauses in phonation.[22]

It is interesting that George Summey, in considering the last quotation, cautions the reader to note the qualification "in large part." The implications in the warning are made explicit by H. A. Gleason's statement that "there is available very little descriptive data on how the English, or any other, punctuation system is actually used. The large volume of published material which is available is predominantly normative and almost wholly based on 'logical' categories."[23]

Nevertheless, attempts to get at the underlying function of our punctuation system are being made by many prominent linguists. W. Nelson Francis feels that while "the punctuation system is primarily a written substitute for intonation in speech . . . it does not represent intonation directly [and] . . . the writer must convert from the intonation system to the punctuation system by way of his understanding of the structural meaning."[24]

Harold Whitehall would place less emphasis on the correspondence of intonation and punctuation. Although he admits that "the traditional purpose of punctuation is to symbolize by means of signs the patterns

heard in speech," he feels that "its most important purpose is 'to make grammar graphic.'"[25] While Francis' and Whitehall's statements are not totally opposed to one another, it is clear that the general attitude of each toward the role of punctuation as a whole must yield a somewhat different analysis of the roles which the individual punctuation symbols play in written language.

George Summey's treatment of the subject is lucid without being oversimplified. While his book, American Punctuation, is principally an inductive study of modern punctuation practice, the conclusions which he draws in his analytical introductory statements (which are to some extent linguistically oriented) place him in a position closer to Francis that to Whitehall (see Summey's statement above and that of Marckwardt, in which he concurs). In his analysis of punctuation's role in writing, Summey argues for the division of the symbols into three main classifications which are characterized by the nature of the symbols that are included in them:

> (1) structural punctuation-the use of marks to indicate paragraph and sentence relations, and grouping within sentences
>
> (2) the use of quotation marks and marks of editorial interpolation and ellipsis, and
>
> (3) the use of certain orthographical or work points--division hyphen, compounding hyphen, apostrophe, and abbreviation period[26]

We find that Francis' classification agrees with that of Summey, for, although the former's classification of punctuation symbols falls into four categories, these categories can be regrouped to coincide with the latter's. Francis groups the symbols as follows:

(1) two morphological marks: ' - [27]

(2) three end marks: . ? !

(3) four internal marks: , ; : --

(4) two special marks: " ([28]

Note that Francis' first classification (morphological marks) coincides with Summey's third (orthographical points); Francis' fourth classification (special marks) coincides with Summey's second (editorial marks); and Francis' second and third classifications (end and internal marks) represent only a more detailed statement of Summey's first classification (structural punctuation) (for the sake of clarity and uniformity, Summey's designations for these three classifications of punctuation--i.e., structural, orthographic, and editorial--will be used hereafter in referring to classifications based on the nature of a punctuation symbol).

In view of the general statements by Perrin, Summey, Marckwardt, Francis--indeed almost every writer on punctuation, that punctuation is connected in some way with syntax and intonation, an impression--however unintentional--is conveyed that "true" punctuation is structural punctuation. Indeed, it would even seem that Summey favors this position, for he states: "The principal subject of the book [American Punctuation] is structural punctuation."[29] It may be that this attitude favoring the treatment of structural punctuation has grown out of the prominence which the difficulties associated with structural punctuation have achieved. Yet, if the difficulties previously touched on with regard to quotation marks and apostrophes are considered, it will immediately be

clear that the classifications of editorial symbols and orthographic symbols require their own full measure of consideration.

The tendency to slight non-structural punctuation is overcome by Harold Whitehall, whose classification of punctuation symbols cuts across those lines set up by Francis and Summey. Whitehall classifies, not on the basis of the nature of the symbols, as does Francis and Summey, but rather on the basis of their function. "As a kind of visual configurational feature of grammar, punctuation cannot be properly understood unless the other grammatical features of the language are also understood."[30] He thus classifies all punctuation as falling into one of the following four categories of function:

a. To link sentences and parts of words.

b. To separate sentences and parts of sentences.

c. To enclose parts of sentences.

d. To indicat omissions.[31]

Whitehall's classification has the disadvantage of implying that all punctuation is essentially of the same nature and that various symbols simply perform different functions in the written language. In addition, he seems to be required to force certain marks of punctuation to conform with his system of classification. For example, in classifying the apostrophe as "omission punctuation," Whitehall states (page 131), "originally, the apostrophe (') indicated the omission of a letter no longer pronounced or deliberately suppressed in pronunciation [cf. The Oxford English Dictionary]. This is what it still indicates when used with the possessive singular forms of nouns . . . the Lord's Prayer (earlier, the Lordes Prayer)." At the same time (page 131) he says of

the apostrophe in the plural possessive (also classified as omission punctuation), "In a purely symbolic function corresponding to nothing in actual speech, it indicates possessive plurals of nouns ..." Clearly, the plural possessive apostrophe cannot be classified as omission punctuation. Even 19th century grammarians objected to this use of the apostrophe because it didn't represent an omitted letter. Moreover, the original function of the apostrophe in forming genitives does not necessarily imply its present function in forming them. In comparing the apostrophe in "Lord's" and "can't," for example, we find that they do not perform the same function and do not mean the same thing, even though at one time they might have. Since Whitehall notes (page 120), "The separating period (.)" (the omission period indicates the omission of several letters, as the "Mr."), one might make the same observation concerning the "omission apostrophe" and the "possessive apostrophe": they are quite distinct in functional use.

But Whitehall's functional classifications are helpful when they are used as sub-classifications of the structural, orthographic, and editorial categories of Francis and Summey. For example, within the structural category we have punctuation used in linking (;), separating (:), and enclosing functions (, . . . ,). Within the editorial category we have punctuation used to indicate omission (. . .) and enclosure (" . . . "_. Within the orthographic category we have punctuation which indicates omission (') and linking (-).

A reasonable system of classification such as the above illustrates might serve as a guide in analyzing the various symbols; it is important to note, however, that certain characteristics of certain marks still

remain unidentified. For example, the orthographic linking mark (-) is non-phonemic,[32] as is the orthographic mark of omission ('). The editorial enclosing marks--(...)--are phonemic, as is the structural separating mark (,). In these situations it is clear that the phonemic characteristic of a mark of punctuation cuts across classificatory lines; yet in spite of this, phonemic value cannot be ignored, for as Francis points out, "There are some points where the two systems [punctuation and intonation] coincide."[33] Thus, another classification of punctuation symbols must be considered: direct intonational representation--or phonemic value.

Among other things, the correct use of certain punctuation marks depends upon the writer's knowledge of certain grammatical features of the language. To illustrate this, Francis[34] points out that a lack of knowledge of grammar can result in a comma being placed between subject and predicate in a sentence in which the subject includes a long restrictive modifier (cf. Murray, p.5). Matters are further complicated in this instance because placement of the comma represents a phonemic analysis of the sentence--a feature of punctuation which Francis believes can be utilized in teaching punctuation. This situation not only illustrates the grammatical basis of certain punctuation, it illustrates the need for the consideration of what, for want of a better name, will be called "zero punctuation"--the ability to leave out punctuation which is called for phonemically but is not required graphically. With regard to such situations, Summey says, "Punctuation is primarily for the reader's eye, with only partial correspondences to the movement, stresses, and inflections of speech. One does not always stop on a punctuation

mark, and there are pauses in oral reading that are not marked by punctuation in written matter. For example:

> He has lived in Cleveland, Ohio, since July, 1946. [No stop at the commas after Cleveland, Ohio, or July. The better date style since July 1946 would represent the oral grouping more clearly.]
>
> Officials said that conditions have changed to such an extent since the original plan was filed that amendment of nearly every paragraph would be necessary before the securities could be approved by the Securities and Exchange Commission. [In oral reading of this sentence, boundaries between breath groups will be marked by at least two brief pauses that are not marked by punctuation.]

Though pauses and inflections in oral reading sometimes give useful hints about punctuation, they are no more a complete guide to punctuation than punctuation marks are to delivery."[35]

Thus, grammatical basis and "zero punctuation" must be added to phonemic value and the categories supplied by Francis, Summey, and Whitehall as classificatory features of punctuation symbols.

Although we have isolated some ten classifications which represent characteristic features of punctuation symbols, it must be admitted that our system of classification, while it throws into relief certain significant features of a symbol, is by no means exhaustive. Moreover, since our classifications are far from exclusive (i.e., there is an alarming degree of overlapping and crisscrossing between categories), we are faced with an apparent complexity such that we may well ask what significance all this analysis has for the teacher of punctuation.

C. Implications For Teaching

Paul Roberts, in his book Patterns of English, has used certain features of his own analysis of the punctuation system in proposing

methods of teaching punctuation which break away from traditional practice. (Traditional practice is here taken to mean the teaching of all punctuation by means of general rules which are based on traditional logical grammar). "You can learn how to punctuate," he says, "just by reading, if you keep your eyes open and notice the structures that are punctuated."[36] This observation of syntactic relations, however, is only a part of his method; a greater portion of his time is spent in exploiting the fact that "certain fetures of the sound system called intonation . . . [are] connected to punctuation habits."[37] His language is couched in linguistic terminology, and he makes extensive use of the contributions which linguistics has made to our knowledge of language. Says Roberts, "We have punctuation in speech as well as in writing. In speech the punctuation is composed of double bar junctures and double cross junctures, plus complicated patterns of pitch and stress."[38] Although he admits that "speech punctuation" and "written punctuation" are not exactly related, he holds that there is a "good deal" of connection which is "worthwhile to observe." As a result of these tenets, Roberts elaborately explains the correspondence of the double bar juncture and the comma; the double cross juncture and the period or semi-colon. (Shades of Lowth!) But he also recognizes the variations of pitch which relate to certain junctures, thereby lending a greater degree of accuracy to his description of the punctuation symbol's relation to intonation.

As with other writers on punctuation, we note that Roberts is primarily concerned with periods, commas, and semi-colons--what has been called "structural punctuation." And it is interesting that he has

utilized in teaching its use the two most prominent features of structural symbols: Their phonemic and grammatical characteristics. In summarizing the application of the grammatical features of a symbol to learning structural punctuation, Roberts says (page 260), "We have seen that we don't ordinarily use commas to separate the parts of noun clusters or verb clusters. We don't usually separate a noun cluster from a verb cluster to which it is tied as its subject. On the other hand, we do mark off certain kinds of sentence modifiers from the patterns they modify." With regard to approaching structural punctuation through an understanding of its phonemic correspondence, he says (pages 260, 261), "We saw that double cross juncture, a fall in pitch, is represented in writing by a period or a semicolon; double bar juncture, a rise in pitch, is represented by a comma; single bar juncture, level pitch across a break in a sentence, is shown by no punctuation."

Though Robert's methods may seem strange or unorthodox, we find that he is not alone in feeling that the phonemic feature of structural punctuation can be effectively used in teaching its use. Francis states (pages 563, 654), "There are some points where the two systems [intonation and punctuation] coincide. It is thus possible to make limited use of intonational clues in teaching punctuation . . . Recognition of the contrasting intonation patterns of restrictive and non-restrictive modifiers and of the intonational signals marking structures such as internal sentence-modifiers and coordinate modifiers not joined by a coordinator can help considerably in the placing of commas." But whether Francis would consider Roberts' use of intonational clues in teaching punctuation as "limited" can only be conjectured. Certainly Whitehall,

who looks upon punctuation as "a kind of visual configurational feature of grammar [which] cannot be properly understood unless the other grammatical features of the language are also understood" would not look upon intonational correspondence as a reliable method of teaching. In fact, his analysis of the punctuation system seems to perpetuate the notion that "punctuation is punctuation" and that all punctuation can be equally well taught by a single method.

The "single method" approach, of course, is the technique employed by most English teachers today in teaching punctuation. No allowances are made for the fact that symbols differ in nature, function, and in many other respects. Indeed, the National Council of Teachers of English in its book The English Language Arts In The Secondary School treats the teaching of capitalization and punctuation under a single heading with no apparent distinction made between capitalization of proper names, the possessive apostrophe, the apostrophe in contractions, separating and enclosing commas, and many other punctuation symbols.[39] The significance of Roberts' approach to teaching punctuation is its strong implication that when other than structural punctuation is taught, different methods will need to be employed. It has previously been noted that certain punctuation marks bear neither phonemic nor grammatical characteristics; it would therefore seem fair to state that the methods described above for teaching the use of periods, commas, and semi-colons are irrelevant as a means of teaching the use of marks not bearing like characteristics.

But even though linguistics has demonstrated that the characteristics of punctuation symbols are significant enough to warrant the

formation of a system of classification, the inference that methods of teaching the use of the symbols should be based on those classifications actually has no foundation. Indeed, it has never been demonstrated that a punctuation symbol's characteristics has any influence whatever on the effectiveness or ineffectiveness of the method which is used to teach the use of that symbol. However, it would seem that the current practice of using a single method of instruction to teach the use of all the punctuation symbols (and sometimes capitalization as well) persists not because of a lack of research which demonstrates the influence of a punctuation symbol's characteristics on the effectiveness of a teaching method, but because of the English teacher's failure to recognize these differentiating characteristics.

FOOTNOTES

Chapter I

[1]Paul Roberts, Patterns of English (New York, 1956), p. 215.

[2]George Summey, American Punctuation (New York, 1949), p. 8.

[3]G. H. Vallins, The Pattern of English (London, 1956), p. 149.

[4]Vallins, p. 149.

[5]Ben Jonson, Grammar (1692), cited by Vallins, p. 151.

[6]Robert Lowth, Short Introduction to English Grammar (1762), cited by Vallins, p. 152.

[7]William Cobbett, A Grammar of the English Language (1817), cited by Vallins, p. 152.

[8]Cited by Vallins, p. 153.

[9]Brenan, cited by Vallins, p. 154.

[10]Cited by Vallins, p. 154.

[11]Lindley Murray, An English Grammar (1795), cited by Vallins, p. 156.

[12]Cited by Vallins, p. 156.

[13]Murray, cited by Vallins, p. 158.

[14]Cited by Vallins, p. 159.

[15]The London Times Literary Supplement, cited by Vallins, p. 160.

[16]James A. H. Murray et al., eds., The Oxford English Dictionary (Oxford, 1933), I, 393.

[17]Cited in The Oxford English Dictionary, I, 393.

[18]Ephraim Chambers, Cyclopedia of Arts and Sciences (1728), cited by Susie I. Tucker, English Examined (Cambridge, England, 1961), p. 75.

[19]James Sledd, A Short Introduction to English Grammar (Chicago, 1959), p. 272.

[20]Porter G. Perrin, Writer's Guide and Index to English (Chicago, 1942), p. 276.

[21]Summey, p. 3.

FOOTNOTES

Chapter I (cont.)

[23]A. H. Gleason, An Introduction to Descriptive Linguistics (New York, 1955), p. 326.

[24]W. Nelson Francis, The Structure of American English (New York, 1958), pp. 563, 564.

[25]Harold Whitehall, Structural Essentials of English (New York, 1951), p. 119.

[26]Summey, p. 3.

[27]Angle brackets are used to indicate a grapheme, which is a group of graphs or symbols belonging to the same family. Thus, A, a, A, and a are allographs of the "a" grapheme, which is written (a).

[28]Francis, p. 472.

[29]Summey, p. 3.

[30]Whitehall, pp. 119, 120.

[31]Whitehall, p. 120.

[32]A phoneme is a class of sounds. For example, the /p/ (just as graphemes are enclosed within angle brackets, phonemes are enclosed within diagonals) in "pin" is easily demonstrated to be different from the /p/'s in "spin" and "drop." These phonetically different /p/'s are all allophones of the same phoneme. English has twenty-four consonant phonemes and nine vowel phonemes. In addition, it has twelve intonational phonemes, four each of stress, pitch, and juncture. A period at the end of a sentence is phonemic because it indicates a certain degree of juncture and drop in pitch. We have said that the linking hyphen (-) is non-phonemic because, unlike the period, it represents no intonational value.

[33]Francis, p. 563.

[34]Francis, p. 476.

[35]Summey, pp. 6,7.

[36]Roberts, p. 217.

[37]Roberts, p. 217.

[38]Roberts, p. 235 (see Chapter 47, p. 227 for definitions and explanations of pitch, stress, and juncture).

[39]The Commission On The English Curriculum of The National Council of Teachers of English, The English Language Arts in the Secondary School (New York, 1956), p. 105.

APPENDIX B

REPORT ON A LITERATURE SEARCH:

STATISTICAL STUDIES OF PUNCTUATION FREQUENCY*

by Mary Lynn Vickers

*Report of a literature search commissioned by Andrew Malcolm and prepared by Mary Lynn Vickers of the Rochester Regional Research Library, 50 West Main Street, Rochester, N.Y.

To: Andrew Malcolm November 12, 1975

From: Mary Lynn Vickers

RE: Report on a Literature Search:

Statistical Studies of Punctuation Frequency

Beginning with Harry Shaw's discussion[1] of punctuation frequency and its gradual decrease during the past century, a wide-ranging literature search was conducted to determine whether any sort of similar study has ever been done. Apparently Mr. Shaw did his own brief enumeration of punctuation markings, for we find no reference to a prior study from this otherwise meticulous foot-noter.

Looking into the New York Times Index from 1974 back to 1968, I found an August 1972 article about a McGraw-Hill editor, A. C. Lloyd, who was producing practice typing sentences using all letters of the alphabet with the fewest possible words.[2] Two articles in 1971 reported on a patent granted for a reading machine able to identify as many as 80 type fonts,[3] and on an American Heritage/Houghton Mifflin study financed to the tune of $3 million.[4] The study was to determine the frequency of words in public and private school books and magazines, and was conducted by P. Davies.[5] None of these leads proved to be relevant. The

---

[1] Shaw, Harry, PUNCTUATE IT RIGHT! New York, Barnes & Noble, 1963. pp. 12-13.

[2] New York Times, August 5, 1972; p. 27, col. 6.

[3] Ibid., February 6, 1971; p. 46, col. 2.

[4] Ibid., September 4, 1971; p. 22, col. 5.

[5] This is the volume we looked at in your office - I don't have the exact reference.

terms searched in the NYT Index were: punctuation, typography, printing, language and language - English.

The next bibliographic source consulted in this search was the Education Index for the period July 1959 to the present (Sept. 1975). Terms searched in this index were: punctuation, type, type frequency and English language - punctuation. A list of possible - sounding articles is attached to this report. (You've been sent the first pages of the first 2 articles on the list.) None of the references however seemed to be exactly on the target topic of this search.

Psychological Abstracts from January 1973 to December 1974 yielded 2 possible abstracts (sent to you already) but still nothing of apparent direct relevance. Terms searched in the Psychological Abstracts were: punctuation, type, letters, morphology (language), linguistics and grammar.

To explore the technology side of printing and machine design for typesetting, Engineering Index was searched from 1966 to 1974 and from 1896 (its beginning date) up to 1910. The recent references included a review of keyboard design (with 83 references), a report on a computer program to hyphenate words, optical character recognition technology, and the proceedings of a computer typesetting conference. Photocopies of these abstracts [were] attached. Terms searched in the recent Engineering Index were: printing, typesetting, and typewriting.

The older references in the 1896-1910 period included "recent advances" in graphic arts and a survey of 50 years of printing technology. Design of a type-casting machine was also mentioned. Photocopies of these possible references are attached to this report. One

additional reference, from V. 2 of the index (not photocopied) is also included.[6] Engineering Index was titled Current Engineering Literature at that time. Terms searched in the older index were: linotype, printing, type-lettering and lithography.

The conclusion seems to be that no comprehensive study of the statistical frequency of punctuation has ever been done, or if it has been done it is certainly not in a useful, widely-used or known format. For people interested in such a study it certainly seems that there is a dearth of relevant information available from the wide variety of sources consulted.

Summary of sources searched:

NY Times Index	1968-1974
Education Index	July 1959 - September 1975
Psychological Abstracts	Jan. 1973 - December 1974
Engineering Index	1966 - 1974
	1896 - 1910

---

[6] Zeithscrift des Vereines Deutscher Ingenieur, Nov. 10, 1894. "Printing Machinery in the U.S." by E. Wentscher. (weekly, pub. in Berlin)

Education Index

(July 1959 - September 1975)

July '70 - June '71	Effects of standard vs. alphabetical formats on keyboard performance. J. Ap. Psychology. R. S. Heirsch. 54:484-90, Dec.'70.
July '63 - June '64	Linguistic characteristics of punctuation symbols and the teaching of punctuation skills. R.S. Zais. Eng. J. 52:677-81, Dec. '63.
July '73 - June '74	In search of an early teaching grammar. J. L. Presland. Educ. Research 16:112-20, Feb.'74.
July '71 - June '72	Usuage and teaching punctuation, no frequency study
Individual Issues July '74 - Sept. '75	Nothing
July '70 - June '71	Quote me on quotation marks. J.T. Leeson. Instr. 80:42, April '71. Psyching out commas: syntactic & semantic relation. J.E. Haney, Col Comp & Comm 21:173-6, May'70.
July '69 - June '70	Nothing
July '68 - June '69	Notes on history of Spanish punctuation. J.R. Chatham. Hispania. 52:79-80, March 1969. Punctuation Personified, R. Hale. Today's Ed. 58:27, March '69.
July '67 - June '68	Nothing
July '65 - June '66	Apostrophe to the Ocean & Heave it in! R.E. Moore, Engl. Jnl. 55:198-200, Feb.'66. Rally round the apostrophe, boys! L.H. Feigenbaum, Hi Points. 48:64-5, Jan. '66.
July '63 - June '64	Capitalization & Punctuation - a diagnostic test. R.R. Odom. Calif. Jnl. of Ed. Research. 15:68-75. March'64.
July '61 - June '63	Punctuating the Compound Sentence. J. C. Gray. Engl. J. 51:573-4, Nov. '62.
July '59 - June '61	Comedy of Commas. M. Nurnberg. High Points. 42:47-50, Apr. '60. Comma no visual handicap. S.C. Lucey. High Points 43:64, Apr. '61. Punctuation and spelling. E. Campbell-Dover. J. Bsns. Ed. 35:215-216. Feb.'60.

APPENDIX C

CLASSIFICATION OF PUNCTUATION MARKS AND THEIR USES

APPENDIX C. PUNCTUATION MARKS AND THEIR USES

INTRODUCTION

Twelve marks of punctuation and 23 subcategories of use were studied in this thesis. These 35 categories are defined herein. They are arranged first alphabetically by the twelve chosen marks of punctuation, and the various uses of these marks are described following the mark. For ease of identification, the abbreviations used in the tables in Appendix D are used here.

APOSTROPHIES: This category includes all apostrophies disclosed in the study and includes, but is not limited to, the subcategories of use which follow. No subcategories have been provided for certain minor uses of the apostrophe, so the sum of the subcategories does not always equal the sum of all apostropies.

POSSESSIVES: This subcategory includes all apostrophies used to form possessives including singular, regular plurals, and irregular plurals, and uninflected words ending in s which are not plurals, e.g., The man's hat was gone; The girls' hats were gone; the women's hats were gone; Jones' hat was gone, respectively.

CONTRACTIONS: This subcategory includes all apostrophies used in contractions, e.g., don't, won't, I've, etc.

COLONS: This category includes all colons disclosed in the study. No subcategories were used because of the low frequency rate of colons.

COMMAS: This category includes all commas disclosed in the study and includes, but is not limited to, the subcategories of use which follow. No subcategories have been provided for certain minor uses

of the comma, so the sum of the subcategories does not always equal the sum of all commas.

SEPARATE QUOTES: This subcategory includes all commas used at either the opening or the closing of a quotation, e.g., Jones said, "I went to the meeting, but I did not stay," when interviewed in New York. In this example, two commas were counted, viz., the one contiguous to the word said and the one contiguous to the word stay; the comma contiguous to the word meeting is not counted in this subcategory (q.v. Independent Clauses).

INTRODUCE PHRASES: This subcategory includes all commas used to separate one or more words at the beginning of a sentence. These take the form of prepositional phrases (e.g., In 1976, the sampling began.), and adverbs which modify clauses (e.g., Occasionally, the President plays golf). Sometimes commas were found to have dual functions where one function was to introduce a phrase, e.g., In Atlanta, the capital of Georgia, the Governor had a press conference. In this example, the comma contiguous to the word Atlanta separates an introductory phrase and also opens an appositive. This would then be counted twice and for this reason, one cannot assume that the sum for the subcategories is equal to the sum of all commas.

SEPARATE SERIES: This subcategory includes all commas used to separate words in series, e.g., The U.S. Flag is red, white, and blue. The comma contiguous to the word red is counted, and the comma contiguous to the word white is counted, when a comma is used in this position. (Modern usage allows the optional use of commas in this position.[1])

This subcategory also includes all commas used to separate phrases in series, e.g., The young girl, her mother, and the elevator operator were stranded. The commas contiguous to the words girl and mother are counted in this subcategory.

INDEPENDENT CLAUSES: This subcategory includes all commas used to separate independent clauses in a sentence, e.g., The attorney for the prosecution spoke to reporters, Smith wanted to speak, but Smith's attorney advised against statements. In this example, the commas contiguous to the words reporters and speak are counted.

ADVERBIAL CLAUSES: This subcategory includes all commas used to separate adverbial clauses from other clauses, e.g., Because he was sick, he had left work early that day. The comma contiguous to the word sick is counted. Sometimes commas were found to have dual functions where one function was to separate an adverbial clause, e.g., Because he had killed Smith, the driver of the other car, Johnson was charged with manslaughter. In this example, the comma contiguous to the word car separates the adverbial clause from the independent clause, and also closes the appositive opened by the comma which is contiguous to the word Smith. This would then be counted twice and for this reason, one cannot assume that the sum of the subcategories is equal to the sum of all commas.

OPEN APPOSITIVES and CLOSE APPOSITIVES: These subcategories include all commas used at the opening and the closing of appositives, e.g., Paul Miller, president of RIT, gave the principal address. In this example, the comma contiguous to the word Miller is counted as an opening appositive, and the comma contiguous to the word RIT is

counted as a closing appositive. Commas used for opening and closing appositives were often found to perform dual functions q.v., INTRODUCE PHRASES and ADVERBIAL CLAUSES.

OPEN PARENTHETICS and CLOSE PARENTHETICS: These subcategories include all commas used at the opening and closing of parenthetic words, phrases, or clauses, e.g., Smith, meanwhile, while trying to fix his car, saw Jones enter the barn, whose door was open, and attempt to steal tools. In this example, the commas contiguous to the words Smith and meanwhile separate the parenthetic word meanwhile, the commas contiguous to the words meanwhile and car separate the parenthetic phrase, 'while trying to fix his car,' and the commas contiguous to the words barn and open separate the parenthetic clause, 'whose door was open,' from the main clause of the example sentence. In this example, the comma contiguous to the word meanwhile performs a dual function and is counted twice. Since this is counted twice, and since commas may perform other concurrent functions, one cannot assume that the sum of the subcategories is equal to the sum of all commas (q.v. INTRODUCE PHRASES and ADVERBIAL CLAUSES.)

DASHES: This category includes all dashes disclosed in the study. No subcategories were used because of the low frequency rate of dashes.

ELLIPSES: This category includes all ellipses disclosed in the study. None were found for any purpose other than omission of words in quotations.

EXCLAMATION POINTS: This category includes all exclamation points disclosed in the study including, but not limited to those in the

subcategory which follows.

(EXCLAMATION POINTS USED) TO TERMINATE: This subcategory includes all exclamation points used to terminate sentences. Although most exclamation points are used for this purpose a few are not (q.v. Tables D-6, D-13, D-23, and D-30 where exclamation points are reported as serving purposes other than terminating sentences). It was vital that this distinction be made so that sentence counts were obtained accurately.

HYPHENS: This category includes all hyphens disclosed in this study and includes, but is not limited to, the subcategories of use which follow. No subcategories have been provided for certain minor uses of the hyphen so the sum of the subcategories does not always equal the sum of all hyphens.

COMPOUND WORDS: This subcategory includes all hyphens used in compound words, e.g. Truman's anti-inflation bill is thought to be pro-labor by many conservatives. In this example, the hyphens in the words anti-inflation and pro-labor are counted as assisting in compounding words. The distinction between compound words and unit modifiers (q.v.) is not always clear and is sometimes historic in nature (e.g., rail road was used in the early part of the nineteenth century, rail-road was used in later part of the nineteenth century and early part of the twentieth century, and modern usage calls for railroad to be composed solid.[2]) When the difference between compound and other uses was in doubt, Chapter 7 of the U.S. Government Printing office Style Manual was used as the standard.[3]

UNIT MODIFIERS: This subcategory includes all hyphens used to join words together into modifying units, e.g., The seven-man board was turned up-side-down by the quick-silver tongue of the member-at-large. All of the hyphens in this example are included as unit modifiers.

WORD BREAKS: This subcategory includes all hyphens used to divide words at the end of a line of type. Sometimes compound words (q.v.) and unit modifiers (q.v.) had hyphens which also serve as word breaks and in such instances, these marks were counted twice. For this reason, one cannot assume that the sum of the subcategories is equal to the sum of all commas.

PAIRS OF PARENTHESES: This category includes all pairs of left and right parentheses disclosed in the study for whatever reason used.

PERIODS: This category includes all periods disclosed in the study and includes, but is not limited to the subcategories of use which follow. No subcategories have been provided for certain minor uses of the period, so the sum of the subcategories does not always equal the sum of all periods.

(PERIODS USED) TO TERMINATE: This subcategory includes all periods used to terminate sentences. It is vital that this distinction be made so that sentence counts were obtained accurately.

ABBREVIATIONS: This subcategory includes all periods used to indicate abbreviations, e.g., The U.S. entered negotiations on Feb. 6 with Gen. Zukov of the U.S.S.R. In this example, all eight periods were used to abbreviate. The period contiguous to the R in U.S.S.R. also serves to terminate the sentence and is counted twice. For

this reason, one cannot assume that the sum of the subcategories is equal to the sum of all periods.

DECIMAL POINTS: This subcategory includes all decimal points, e.g., The new Ford automobile is offered with a standard 1.4-liter engine or a 1.6-liter engine for an extra $87.15. In this example, three decimal points are used. (A period is also used to terminate this sentence.)

QUESTION MARKS: This category includes all question marks disclosed in the study and includes, but is not limited to, the subcategories of use which follow. No subcategories have been provided for certain minor uses of the question mark, so the sum of the subcategories does not always equal the sum of all question marks.

(QUESTION MARKS used) TO TERMINATE: This subcategory includes all question marks used to terminate sentences. Although most question marks are used for this purpose, a few are not (q.v. Tables D-25 and D-30 where question marks are reported as serving purposes other than terminating sentences.) It was vital that this distinction be made so that sentence counts were obtained accurately.

QUOTATION MARKS: This category includes all quotation marks including both double and single, opening and closing, and for any purpose whatsoever including double marks used to open and close direct quotations that were disclosed in this study. Since no subcategories are provided for any use other than opening and closing direct quotations, the sum of these subcategories frequently does not equal the sum of all quotation marks.

OPEN QUOTATIONS and CLOSE QUOTATIONS: These subcategories include all double opening and all double closing quotation marks used to set-off direct quotations, e.g., "The defendant said, 'I killed him,' when I approached," said today's witness for the prosecution in this "kangaroo court." In this example, only one opening and one closing mark are reported for this sentence. The single quotation marks used within the major quote and the double quotation marks used to include the words kangaroo court are reported in the total QUOTATION MARKS category.

SEMICOLONS: This category includes all semicolons disclosed in this study and includes, but is not limited to, the subcategories of use which follow. No subcategories have been provided for certain minor uses of the semicolon, so the sum of the subcategories does not always equal the sum of all semicolons.

SEPARATE IND. (Independent) CLAUSES: This subcategory includes all semicolons used to separate independent clauses, e.g., The students collected on the beach; they watched the police arrive to disperse them; they became an angry mob. In this example, the two semicolons would be counted in this subcategory.

SUPERIOR COMMAS: This subcategory includes all semicolons used as superior commas, e.g., The International Alphabet Flags have red, white and blue horizontal stripes for the letter C; red, white and blue vertical stripes for the letter T; and a black, yellow, red, and blue triangle for the letter Z. In this example, there are two semicolons used as superior commas contiguous to the letters C and T.

APPENDIX D

TABLES OF PUNCTUATION MARKS BY MARK AND USE OF MARK

TABLE D - 1

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM PAGE 1 OF THE NEW YORK TIMES FOR FEBRUARY 29, 1970 (N = 300)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPHES	3.33333	57.7350	196.000
POSSESSIVES	3.33333	57.7350	196.000
CONTRACTIONS	0	0	0
COLONS	0	0	0
COMMAS	53.3333	225.073	41.7551
SEPARATE QUOTES	0	0	0
INTRODUCE PHRASES	0	0	0
SEPARATE SERIES	6.66667	81.5130	138.301
INDEPENDENT CLAUSES	0	0	0
ADVERBIAL CLAUSES	0	0	0
OPEN APPOSITIVES	13.3333	114.889	97.5071
CLOSE APPOSITIVES	0	0	0
OPEN PARENTHETICS	33.3333	179.305	61.0407
CLOSE PARENTHETICS	0	0	0
DASHES	10.0000	99.6650	112.702
ELLIPSES	0	0	0
EXCLAMATION POINTS	0	0	0
TO TERMINATE	0	0	0
HYPHENS	43.3333	203.947	53.2586
COMPOUND WORDS	3.33333	57.7350	196.000
UNIT MODIFIERS	0	0	0
WORD BREAKS	43.3333	203.947	53.2586
PAIRS OF PARENTHESES	0	0	0
PERIODS	46.6667	211.276	51.2391
TO TERMINATE	23.3333	151.214	73.3340
ABBREVIATIONS	16.6667	128.233	87.6650
DECIMAL POINTS	3.33333	57.7350	196.000
QUESTION MARKS	0	0	0
TO TERMINATE	0	0	0
QUOTATION MARKS	13.3333	114.889	97.5071
OPEN QUOTATIONS	0	0	0
CLOSE QUOTES	0	0	0
SEMICOLONS	3.33333	57.7350	196.000
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	3.33333	57.7350	196.000

TABLE D - 2

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM
PAGE 1 OF THE LOS ANGELES TIMES
FOR FEBRUARY 29, 1976 (N = 300)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPIES	16.6667	120.233	87.0656
POSSESSIVES	16.6667	120.233	87.0656
CONTRACTIONS	0	0	0
COLONS	0	0	0
COMMAS	26.6667	161.376	69.4805
SEPARATE QUOTES	3.33333	57.7350	196.000
INTRODUCE PHRASES	3.33333	57.7350	196.000
SEPARATE SERIES	0	0	0
INDEPENDENT CLAUSES	0	0	0
ADVERBIAL CLAUSES	3.33333	57.7350	196.000
OPEN APPOSITIVES	6.66667	81.5120	120.361
CLOSE APPOSITIVES	0	0	0
OPEN PARENTHETICS	13.3333	114.889	97.5071
CLOSE PARENTHETICS	0	0	0
DASHES	0	0	0
ELLIPSES	0	0	0
EXCLAMATION POINTS	0	0	0
TO TERMINATE	0	0	0
HYPHENS	56.6667	231.521	40.8476
COMPOUND WORDS	0	0	0
UNIT MODIFIERS	13.3333	114.089	97.5071
WORD BREAKS	43.3333	203.047	52.2586
PAIRS OF PARENTHESES	3.33333	57.7350	196.000
PERIODS	66.6667	240.961	42.4116
TO TERMINATE	40.0000	196.257	55.5298
ABBREVIATIONS	26.6667	161.376	60.4805
DECIMAL POINTS	0	0	0
QUESTION MARKS	0	0	0
TO TERMINATE	0	0	0
QUOTATION MARKS	33.3333	179.905	61.0407
OPEN QUOTATIONS	20.0000	140.234	79.3448
CLOSE QUOTES	13.3333	114.089	97.5071
SEMICOLONS	0	0	0
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	0	0	0

TABLE D - 3

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM
PAGE 1 OF THE CHICAGO TRIBUNE
FOR FEBRUARY 29, 1976 (N = 300)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPIES	6.66667	81.5130	138.361
POSSESSIVES	6.66667	81.5130	138.361
CONTRACTIONS	0	0	0
COLONS	0	0	0
COMMAS	33.3333	179.805	61.0407
SEPARATE QUOTES	10.0000	99.6650	112.782
INTRODUCE PHRASES	6.66667	81.5130	138.361
SEPARATE SERIES	0	0	0
INDEPENDENT CLAUSES	0	0	0
ADVERBIAL CLAUSES	0	0	0
OPEN APPOSITIVES	6.66667	81.5130	138.361
CLOSE APPOSITIVES	0	0	0
OPEN PARENTHETICS	6.66667	81.5130	138.361
CLOSE PARENTHETICS	0	0	0
DASHES	3.33333	57.7350	196.000
ELLIPSES	0	0	0
EXCLAMATION POINTS	0	0	0
TO TERMINATE	0	0	0
HYPHENS	46.6667	211.270	51.2318
COMPOUND WORDS	33.3333	179.805	61.0407
UNIT MODIFIERS	0	0	0
WORD BREAKS	13.3333	114.989	97.5071
PAIRS OF PARENTHESES	0	0	0
PERIODS	40.0000	196.287	55.5298
TO TERMINATE	33.3333	179.805	61.0407
ABBREVIATIONS	6.66667	81.5130	138.361
DECIMAL POINTS	0	0	0
QUESTION MARKS	0	0	0
TO TERMINATE	0	0	0
QUOTATION MARKS	23.3333	151.212	73.3340
OPEN QUOTATIONS	10.0000	99.6650	112.782
CLOSE QUOTES	10.0000	99.6650	112.782
SEMICOLONS	0	0	0
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	0	0	0

TABLE D - 4

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM PAGE 1 OF THE ATLANTA JOURNAL AND CONSTITUTION FOR FEBRUARY 29, 1976 (N = 300)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPIES	10.0000	99.6650	112.782
POSSESSIVES	3.33333	57.7350	196.000
CONTRACTIONS	6.66667	81.5130	138.361
COLONS	0	0	0
COMMAS	53.3333	225.073	47.7551
SEPARATE QUOTES	3.33333	57.7350	196.000
INTRODUCE PHRASES	16.6667	128.233	87.0650
SEPARATE SERIES	0	0	0
INDEPENDENT CLAUSES	10.0000	99.6650	112.782
ADVERBIAL CLAUSES	0	0	0
OPEN APPOSITIVES	3.33333	57.7350	196.000
CLOSE APPOSITIVES	0	0	0
OPEN PARENTHETICS	13.3333	114.880	97.5071
CLOSE PARENTHETICS	0	0	0
DASHES	0	0	0
ELLIPSES	0	0	0
EXCLAMATION POINTS	0	0	0
TO TERMINATE	0	0	0
HYPHENS	50.0000	237.984	44.8851
COMPOUND WORDS	0	0	0
UNIT MODIFIERS	3.33333	57.7350	196.000
WORD BREAKS	53.3333	225.073	47.7551
PAIRS OF PARENTHESES	0	0	0
PERIODS	46.6667	211.276	51.2318
TO TERMINATE	43.3333	203.947	53.2586
ABBREVIATIONS	3.33333	57.7350	196.000
DECIMAL POINTS	0	0	0
QUESTION MARKS	0	0	0
TO TERMINATE	0	0	0
QUOTATION MARKS	6.66667	81.5130	138.361
OPEN QUOTATIONS	3.33333	57.7350	196.000
CLOSE QUOTES	3.33333	57.7350	196.000
SEMICOLONS	0	0	0
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	0	0	0

TABLE D - 5

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM
PAGE 1 OF THE DENVER POST
FOR FEBRUARY 29, 1976 (N = 300)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPIES	20.0000	140.234	79.3448
POSSESSIVES	16.6667	128.233	87.0656
CONTRACTIONS	3.33333	57.7350	196.000
COLONS	0	0	0
COMMAS	40.0000	196.297	55.5298
SEPARATE QUOTES	0	0	0
INTRODUCE PHRASES	3.33333	57.7350	196.000
SEPARATE SERIES	13.3333	114.889	97.5071
INDEPENDENT CLAUSES	3.33333	57.7350	196.000
ADVERBIAL CLAUSES	0	0	0
OPEN APPOSITIVES	0	0	0
CLOSE APPOSITIVES	0	0	0
OPEN PARENTHETICS	13.3333	114.889	97.5071
CLOSE PARENTHETICS	0	0	0
DASHES	6.66667	81.5130	136.301
ELLIPSES	0	0	0
EXCLAMATION POINTS	0	0	0
TO TERMINATE	0	0	0
HYPHENS	36.6667	188.256	58.0996
COMPOUND WORDS	3.33333	57.7350	196.000
UNIT MODIFIERS	0	0	0
WORD BREAKS	33.3333	179.805	61.0407
PAIRS OF PARENTHESES	3.33333	57.7350	196.000
PERIODS	60.0000	237.884	44.8651
TO TERMINATE	30.6667	188.256	58.0996
ABBREVIATIONS	0	0	0
DECIMAL POINTS	23.3333	151.212	73.3340
QUESTION MARKS	0	0	0
TO TERMINATE	0	0	0
QUOTATION MARKS	0	0	0
OPEN QUOTATIONS	0	0	0
CLOSE QUOTES	0	0	0
SEMICOLONS	0	0	0
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	0	0	0

TABLE D - 6

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM
PAGE 1 OF THE NEW YORK TIMES
FOR FEBRUARY 29, 1970 (N = 2135)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPIES	8.009370	93.9374	44.7754
POSSESSIVES	8.009370	93.9374	44.7754
CONTRACTIONS	0	0	0
COLONS	1.40515	37.4676	113.106
COMMAS	56.6745	231.274	17.3170
SEPARATE QUOTES	1.87354	43.2539	97.9311
INTRODUCE PHRASES	6.55738	80.7306	52.2234
SEPARATE SERIES	6.55738	80.7306	52.2234
INDEPENDENT CLAUSES	.936768	30.5995	130.500
ADVERBIAL CLAUSES	1.87354	43.2539	97.9311
OPEN APPOSITIVES	6.55738	80.7306	52.2234
CLOSE APPOSITIVES	.468384	21.6422	196.000
OPEN PARENTHETICS	15.4567	123.289	33.8024
CLOSE PARENTHETICS	9.36768	96.3550	42.6314
DASHES	4.21546	64.9047	65.2108
ELLIPSES	0	0	0
EXCLAMATION POINTS	.468384	21.6422	196.000
TO TERMINATE	0	0	0
HYPHENS	52.9274	229.941	17.9478
COMPOUND WORDS	1.40515	37.4676	113.108
UNIT MODIFIERS	6.08899	77.8123	54.2076
WORD BREAKS	43.0913	208.110	19.9940
PAIRS OF PARENTHESES	0	0	0
PERIODS	53.3958	224.874	17.8641
TO TERMINATE	34.6604	182.961	22.3914
ABBREVIATIONS	17.7986	132.250	31.5186
DECIMAL POINTS	.468384	21.6422	196.000
QUESTION MARKS	0	0	0
TO TERMINATE	0	0	0
QUOTATION MARKS	16.8618	135.866	34.1794
OPEN QUOTATIONS	7.96253	88.8979	47.3504
CLOSE QUOTES	7.49415	86.2639	48.8275
SEMICOLONS	0	0	0
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	0	0	0

TABLE D - 7

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM
PAGE 1 OF THE LOS ANGELES TIMES
FOR FEBRUARY 29, 1976 (N = 2201)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPIES	9.99546	99.4990	41.58 74
POSSESSIVES	8.63244	92.5301	44.7012
CONTRACTIONS	1.81736	42.6014	97.9332
COLONS	1.36302	36.9023	113.100
COMMAS	42.7079	202.244	19.7640
SEPARATE QUOTES	4.54339	67.2667	61.8537
INTRODUCE PHRASES	4.99773	70.5338	58.9616
SEPARATE SERIES	3.63471	60.1926	69.1861
INDEPENDENT CLAUSES	1.81736	42.6014	97.9332
ADVERBIAL CLAUSES	2.72603	52.1521	79.9257
OPEN APPOSITIVES	1.08905	63.8293	65.2144
CLOSE APPOSITIVES	0	0	0
OPEN PARENTHETICS	21.8083	146.090	27.9063
CLOSE PARENTHETICS	1.08905	63.8293	65.2144
DASHES	4.54339	67.2667	61.8537
ELLIPSES	0	0	0
EXCLAMATION POINTS	0	0	0
TO TERMINATE	0	0	0
HYPHENS	51.7946	233.643	18.8457
COMPOUND WORDS	4.54339	67.2667	61.8537
UNIT MODIFIERS	7.72376	87.5046	47.3630
WORD BREAKS	38.9188	192.720	20.9404
PAIRS OF PARENTHESES	2.72603	87.3648	103.240
PERIODS	51.7946	221.663	17.8794
TO TERMINATE	35.4384	184.927	21.8008
ABBREVIATIONS	9.54112	97.2336	42.5759
DECIMAL POINTS	6.81508	82.2904	50.4457
QUESTION MARKS	0	0	0
TO TERMINATE	0	0	0
QUOTATION MARKS	16.3562	130.404	33.3083
OPEN QUOTATIONS	8.17810	90.0828	46.0188
CLOSE QUOTES	6.81508	82.2904	50.4457
SEMICOLONS	1.36302	36.9023	113.100
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	0	0	0

TABLE D - 8

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM
PAGE 1 OF THE CHICAGO TRIBUNE
FOR FEBRUARY 29, 1976 (N = 2107)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPIES	15.6621	121.194	33.9530
POSSESSIVES	11.3906	100.142	39.7693
CONTRACTIONS	1.89843	43.5400	97.9902
COLONS	.949217	30.9020	138.560
COMMAS	55.0546	228.141	17.3943
SEPARATE QUOTES	6.16991	78.3247	54.2055
INTRODUCE PHRASES	6.16991	78.3247	54.2055
SEPARATE SERIES	4.74608	68.7445	61.8481
INDEPENDENT CLAUSES	1.42382	37.7157	113.107
ADVERBIAL CLAUSES	1.89843	43.5400	97.9902
OPEN APPOSITIVES	5.69530	75.2699	50.4324
CLOSE APPOSITIVES	2.84765	53.3000	79.9216
OPEN PARENTHETICS	18.0351	133.110	31.5149
CLOSE PARENTHETICS	9.49217	96.9873	43.6288
DASHES	2.37304	48.6670	87.5706
ELLIPSES	0	0	0
EXCLAMATION POINTS	.474608	21.7855	196.000
TO TERMINATE	.474608	21.7855	196.000
HYPHENS	43.6640	204.395	19.9881
COMPOUND WORDS	.949217	30.9020	138.560
UNIT MODIFIERS	4.74608	68.7445	61.8481
WORD BREAKS	40.9163	197.911	20.7043
PAIRS OF PARENTHESES	0	0	0
PERIODS	63.5975	244.093	16.3095
TO TERMINATE	48.4101	214.682	18.9358
ABBREVIATIONS	18.9843	136.502	30.7020
DECIMAL POINTS	.949217	30.9020	138.560
QUESTION MARKS	0	0	0
TO TERMINATE	0	0	0
QUOTATION MARKS	25.6289	161.039	26.8303
OPEN QUOTATIONS	9.96678	99.3586	42.5671
CLOSE QUOTES	9.49217	96.9873	43.6288
SEMICOLONS	1.42303	37.7157	113.107
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	0	0	0

TABLE D - 9

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM PAGE 1 OF THE ATLANTA JOURNAL AND CONSTITUTION FOR FEBRUARY 29, 1976 (N = 1649)

MARK & USE	MEAN/A-WORD	STD.DEV.	+/- %
APOSTROPIES	11.5221	106.753	44.7192
POSSESSIVES	5.45785	73.6978	65.1745
CONTRACTIONS	7.98357	88.4656	54.1623
COLONS	.606428	24.6258	196
COMMAS	49.7271	217.447	21.16[illegible]
SEPARATE QUOTES	1.81928	42.6272	113.002
INTRODUCE PHRASES	10.3003	101.041	47.3057
SEPARATE SERIES	9.70285	98.0537	48.7705
INDEPENDENT CLAUSES	10.3003	101.041	47.3057
ADVERBIAL CLAUSES	1.81928	42.6272	113.002
OPEN APPOSITIVES	2.42571	49.2467	97.9108
CLOSE APPOSITIVES	1.81928	42.6272	113.002
OPEN PARENTHETICS	10.3003	101.041	47.3057
CLOSE PARENTHETICS	3.63857	60.2290	79.8952
DASHES	2.42571	49.2467	97.9108
ELLIPSES	0	0	0
EXCLAMATION POINTS	0	0	0
TO TERMINATE	0	0	0
HYPHENS	42.4500	201.675	22.9308
COMPOUND WORDS	0	0	0
UNIT MODIFIERS	7.27714	85.0210	56.3912
WORD BREAKS	34.5664	182.734	25.5159
PAIRS OF PARENTHESES	1.21286	34.8155	138.551
PERIODS	62.4621	242.066	18.7053
TO TERMINATE	52.7592	223.620	20.4578
ABBREVIATIONS	8.48999	91.7770	52.1702
DECIMAL POINTS	3.03214	54.9980	87.5474
QUESTION MARKS	.606428	24.6258	196
TO TERMINATE	.606428	24.6258	196
QUOTATION MARKS	12.7350	122.506	46.4305
OPEN QUOTATIONS	3.63857	60.2290	79.8952
CLOSE QUOTES	3.03214	54.9980	87.5474
SEMICOLONS	0	0	0
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	0	0	0

TABLE D - 1W

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM
PAGE 1 OF THE DENVER POST
FOR FEBRUARY 29, 1970 (N = 2671)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPIES	14.9757	121.478	30.7032
POSSESSIVES	7.48783	86.2238	43.6707
CONTRACTIONS	6.36866	79.5394	47.3943
COLONS	2.24635	47.3513	79.9417
COMMAS	53.5380	225.146	15.9485
SEPARATE QUOTES	3.36952	57.9606	65.2354
INTRODUCE PHRASES	9.35979	96.3102	39.0234
SEPARATE SERIES	8.23662	90.3902	41.6227
INDEPENDENT CLAUSES	1.87196	43.2337	87.5882
ADVERBIAL CLAUSES	1.12317	33.5012	113.118
OPEN APPOSITIVES	3.36952	57.9606	65.2354
CLOSE APPOSITIVES	1.87196	43.2337	87.5882
OPEN PARENTHETICS	14.2269	118.447	31.5743
CLOSE PARENTHETICS	7.86222	88.3305	42.6102
DASHES	3.74392	61.0843	61.8701
ELLIPSES	0	0	0
EXCLAMATION POINTS	0	0	0
TO TERMINATE	0	0	0
HYPHENS	30.6904	188.036	19.4360
COMPOUND WORDS	1.49757	38.6767	97.9449
UNIT MODIFIERS	7.11344	84.0504	44.8137
WORD BREAKS	20.9562	161.986	22.7096
PAIRS OF PARENTHESES	1.87196	43.2337	87.5882
PERIODS	54.6812	227.360	15.7745
TO TERMINATE	39.3111	194.371	18.7514
ABBREVIATIONS	11.2317	105.403	35.5897
DECIMAL POINTS	4.11831	64.0538	56.9855
QUESTION MARKS	1.12317	33.5012	113.118
TO TERMINATE	1.12317	33.5012	113.118
QUOTATION MARKS	24.3355	166.067	26.1916
OPEN QUOTATIONS	9.73418	98.1990	38.2504
CLOSE QUOTES	9.35979	96.3102	39.0234
SEMICOLONS	0	0	0
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	0	0	0

TABLE D - 11

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM PAGE 2 OF THE NEW YORK TIMES FOR FEBRUARY 29, 1976 (N - 797)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPIES	11.2923	105.730	65.0042
POSSESSIVES	11.2923	105.730	65.0042
CONTRACTIONS	0	0	0
COLONS	2.50941	50.0625	138.506
COMMAS	46.4241	210.534	31.4851
SEPARATE QUOTES	2.50941	50.0625	138.506
INTRODUCE PHRASES	0	0	0
SEPARATE SERIES	0	0	0
INDEPENDENT CLAUSES	1.25471	35.4218	196.000
ADVERBIAL CLAUSES	2.50941	50.0625	138.506
OPEN APPOSITIVES	13.8018	116.741	58.7236
CLOSE APPOSITIVES	6.27353	79.2063	87.4334
OPEN PARENTHETICS	7.52823	86.4924	79.7650
CLOSE PARENTHETICS	3.76412	61.2753	113.018
DASHES	2.50941	50.0625	138.506
ELLIPSES	0	0	0
EXCLAMATION POINTS	0	0	0
TO TERMINATE	0	0	0
HYPHENS	40.1506	202.730	35.0553
COMPOUND WORDS	1.25471	35.4218	196.000
UNIT MODIFIERS	5.01882	70.7100	97.8152
WORD BREAKS	27.6035	163.937	41.2324
PAIRS OF PARENTHESES	0	0	0
PERIODS	71.5182	257.850	25.0310
TO TERMINATE	47.6788	213.220	31.0476
ABBREVIATIONS	20.0753	140.346	48.5361
DECIMAL POINTS	2.50941	50.0625	138.506
QUESTION MARKS	0	0	0
TO TERMINATE	0	0	0
QUOTATION MARKS	26.3486	204.984	54.0115
OPEN QUOTATIONS	6.27353	79.0063	87.4334
CLOSE QUOTES	7.52823	86.4924	79.7650
SEMICOLONS	0	0	0
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	0	0	0

TABLE D - 12

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM PAGE 2 OF THE LOS ANGELES TIMES FOR FEBRUARY 29, 1976 (n = 923)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPIES	7.58397	36.9021	73.8396
POSSESSIVES	7.58397	36.9021	73.8396
CONTRACTIONS	0	0	0
COLONS	0	0	0
COMMAS	42.2535	201.276	30.7315
SEPARATE QUOTES	1.08342	32.9154	196.000
INTRODUCE PHRASES	1.08342	32.9154	196.000
SEPARATE SERIES	3.25027	56.9493	113.038
INDEPENDENT CLAUSES	4.33369	65.7236	97.8404
ADVERBIAL CLAUSES	1.08342	32.9154	196.000
OPEN APPOSITIVES	1.08342	32.9154	196.000
CLOSE APPOSITIVES	1.08342	32.9154	196.000
OPEN PARENTHETICS	17.3346	130.586	48.5998
CLOSE PARENTHETICS	8.66739	92.7447	69.0329
DASHES	0	0	0
ELLIPSES	1.08342	32.9154	196.000
EXCLAMATION POINTS	0	0	0
TO TERMINATE	0	0	0
HYPHENS	58.5049	234.823	25.8943
COMPOUND WORDS	4.33369	65.7236	97.8404
UNIT MODIFIERS	3.25027	56.9493	113.038
WORD BREAKS	52.0043	222.156	27.5597
PAIRS OF PARENTHESES	0	0	0
PERIODS	52.0043	222.156	27.5597
TO TERMINATE	39.0033	193.708	32.0006
ABBREVIATIONS	13.0011	113.340	50.2418
DECIMAL POINTS	0	0	0
QUESTION MARKS	0	0	0
TO TERMINATE	0	0	0
QUOTATION MARKS	8.66739	92.7447	69.0329
OPEN QUOTATIONS	4.33369	65.7236	97.8404
CLOSE QUOTES	4.33369	65.7236	97.8404
SEMICOLONS	0	0	0
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	0	0	0

TABLE D - 13

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM PAGE 2 OF THE CHICAGO TRIBUNE FOR FEBRUARY 29, 1976 (N = 966)

MARK & USE	MEAN/K-WORD	STD. DEV.	+/- %
APOSTROPIES	10.3926	101.472	85.0305
POSSESSIVES	5.77367	75.8088	87.4510
CONTRACTIONS	4.61894	67.8448	97.8299
COLONS	0	0	0
COMMAS	86.6051	281.418	21.6424
SEPARATE QUOTES	11.5473	106.898	61.6574
INTRODUCE PHRASES	11.5473	106.898	61.6574
SEPARATE SERIES	0	0	0
INDEPENDENT CLAUSES	5.77367	75.8088	87.4510
ADVERBIAL CLAUSES	0	0	0
OPEN APPOSITIVES	4.61894	67.8449	97.8299
CLOSE APPOSITIVES	3.46420	58.7894	113.030
OPEN PARENTHETICS	34.6420	182.977	35.1796
CLOSE PARENTHETICS	17.3210	130.540	50.1958
DASHES	2.30947	48.0291	136.513
ELLIPSES	0	0	0
EXCLAMATION POINTS	1.15473	33.9814	196.000
TO TERMINATE	0	0	0
HYPHENS	39.2610	194.367	32.9663
COMPOUND WORDS	1.15473	33.9814	196.000
UNIT MODIFIERS	4.61894	67.8448	97.8299
WORD BREAKS	35.7968	185.890	34.5868
PAIRS OF PARENTHESES	0	0	0
PERIODS	69.2841	254.083	24.4253
TO TERMINATE	47.3441	212.497	29.8939
ABBREVIATIONS	21.9400	146.572	44.4952
DECIMAL POINTS	1.15473	33.9814	196.000
QUESTION MARKS	0	0	0
TO TERMINATE	0	0	0
QUOTATION MARKS	31.1778	198.719	42.4513
OPEN QUOTATIONS	16.1663	126.188	51.9881
CLOSE QUOTES	12.7021	112.050	58.7536
SEMICOLONS	0	0	0
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	0	0	0

TABLE D - 14

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM PAGE 2 OF THE ATLANTA JOURNAL AND CONSTITUTION FOR FEBRUARY 29, 1976 (N = 1353)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPIES	13.3038	114.615	45.9063
POSSESSIVES	11.9256	108.141	48.7274
CONTRACTIONS	1.47820	38.4331	138.542
COLONS	1.47820	38.4331	138.542
COMMAS	66.5188	252.229	20.2049
SEPARATE QUOTES	5.17369	71.7086	73.9165
INTRODUCE PHRASES	11.9256	108.141	48.7274
SEPARATE SERIES	13.3038	114.615	45.9063
INDEPENDENT CLAUSES	4.43459	66.4695	79.9690
ADVERBIAL CLAUSES	2.21729	47.0533	113.077
OPEN APPOSITIVES	6.65188	81.3174	65.1398
CLOSE APPOSITIVES	3.69549	60.7906	87.5241
OPEN PARENTHETICS	10.3474	101.232	52.1308
CLOSE PARENTHETICS	2.95639	54.3124	97.8912
DASHES	2.95639	54.3124	97.8912
ELLIPSES	0	0	0
EXCLAMATION POINTS	0	0	0
TO TERMINATE	0	0	0
HYPHENS	38.4331	192.311	26.6627
COMPOUND WORDS	0	0	0
UNIT MODIFIERS	.739008	27.1864	196.000
WORD BREAKS	38.4331	192.311	26.6627
PAIRS OF PARENTHESES	.739008	27.1864	196.000
PERIODS	55.4324	228.907	22.0040
TO TERMINATE	39.9113	195.823	28.1442
ABBREVIATIONS	13.3038	114.615	45.9063
DECIMAL POINTS	2.95639	54.3124	97.8912
QUESTION MARKS	0	0	0
TO TERMINATE	0	0	0
QUOTATION MARKS	25.1293	156.576	33.2010
OPEN QUOTATIONS	11.9256	108.141	48.7274
CLOSE QUOTES	10.3474	101.232	52.1308
SEMICOLONS	0	0	0
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	0	0	0

TABLE D - 15

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM
PAGE 2 OF THE DENVER POST
FOR FEBRUARY 29, 1970 (N = 341)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPIES	17.5953	131.668	79.4261
POSSESSIVES	11.7302	107.827	97.5067
CONTRACTIONS	5.86510	76.4713	138.389
COLONS	11.7302	107.827	97.5607
COMMAS	123.167	349.112	28.3614
SEPARATE QUOTES	2.93255	54.1530	196.000
INTRODUCE PHRASES	2.93255	54.1530	196.000
SEPARATE SERIES	76.2463	265.782	36.0986
INDEPENDENT CLAUSES	8.79765	93.5195	112.827
ADVERBIAL CLAUSES	8.79765	93.5195	112.827
OPEN APPOSITIVES	0	0	0
CLOSE APPOSITIVES	0	0	0
OPEN PARENTHETICS	14.6628	120.375	87.1367
CLOSE PARENTHETICS	5.86510	76.4713	138.389
DASHES	0	0	0
ELLIPSES	0	0	0
EXCLAMATION POINTS	0	0	0
TO TERMINATE	0	0	0
HYPHENS	32.2581	192.851	63.4546
COMPOUND WORDS	2.93255	54.1530	196.000
UNIT MODIFIERS	8.79765	93.5195	112.827
WORD BREAKS	14.6628	120.375	87.1367
PAIRS OF PARENTHESES	0	0	0
PERIODS	55.7185	229.714	43.7590
TO TERMINATE	52.7859	223.934	45.0279
ABBREVIATIONS	2.93255	54.1530	196.000
DECIMAL POINTS	0	0	0
QUESTION MARKS	2.93255	54.1530	196.000
TO TERMINATE	2.93255	54.1530	196.000
QUOTATION MARKS	17.5953	131.668	79.4261
OPEN QUOTATIONS	5.86510	76.4713	138.389
CLOSE QUOTES	5.86510	76.4713	138.389
SEMICOLONS	0	0	0
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	0	0	0

TABLE D - 16

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM
PAGE 1 OF THE NEW YORK TIMES
FOR FEBRUARY 27, 1966 (N = 2377)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPIES	10.0968	99.9951	39.8142
POSSESSIVES	8.83467	93.5965	42.5903
CONTRACTIONS	1.26210	35.5110	113.113
COLONS	.841307	29.0007	136.504
COMMAS	49.2217	216.376	17.6723
SEPARATE QUOTES	2.94489	54.1983	73.9874
INTRODUCE PHRASES	3.78629	61.4291	65.2233
SEPARATE SERIES	5.04838	70.9873	56.4492
INDEPENDENT CLAUSES	1.68279	40.9960	97.9381
ADVERBIAL CLAUSES	1.68279	40.9960	97.9381
OPEN APPOSITIVES	6.73117	81.7843	48.9451
CLOSE APPOSITIVES	3.36559	57.9282	69.1943
OPEN PARENTHETICS	13.8830	117.030	33.8887
CLOSE PARENTHETICS	7.15187	84.2835	47.3707
DASHES	0	0	0
ELLIPSES	0	0	0
EXCLAMATION POINTS	0	0	0
TO TERMINATE	0	0	0
HYPHENS	58.4771	234.693	16.1345
COMPOUND WORDS	0	0	0
UNIT MODIFIERS	7.15187	84.2835	47.3707
WORD BREAKS	50.4838	218.987	17.4385
PAIRS OF PARENTHESES	0	0	0
PERIODS	60.1599	237.833	15.8930
TO TERMINATE	39.5456	194.930	19.8163
ABBREVIATIONS	19.7728	139.248	28.3114
DECIMAL POINTS	.420698	20.5109	196.000
QUESTION MARKS	0	0	0
TO TERMINATE	0	0	0
QUOTATION MARKS	25.2419	156.892	24.9674
OPEN QUOTATIONS	12.2003	109.882	36.1612
CLOSE QUOTES	11.3589	105.993	37.5133
SEMICOLONS	1.26210	35.5110	113.113
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	1.26210	35.5110	113.113

TABLE D - 17

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM PAGE 1 OF THE LOS ANGELES TIMES FOR FEBRUARY 27, 1966 (N = 2273)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPIES	6.59921	80.9848	50.4509
POSSESSIVES	6.59921	80.9848	50.4509
CONTRACTIONS	0	0	0
COLONS	5.71931	75.4262	54.2169
COMMAS	41.7950	200.164	19.6888
SEPARATE QUOTES	.439947	20.9749	196.000
INTRODUCE PHRASES	5.27937	72.4831	56.4432
SEPARATE SERIES	3.51958	59.2346	69.1896
INDEPENDENT CLAUSES	1.75979	41.9221	97.9353
ADVERBIAL CLAUSES	1.75979	41.9221	97.9353
OPEN APPOSITIVES	3.95952	62.8139	65.2102
CLOSE APPOSITIVES	3.95952	62.8139	65.2102
OPEN PARENTHETICS	12.7585	112.255	36.1713
CLOSE PARENTHETICS	6.59921	80.9848	50.4508
DASHES	3.51958	59.2346	69.1896
ELLIPSES	0	0	0
EXCLAMATION POINTS	0	0	0
TO TERMINATE	0	0	0
HYPHENS	57.1931	236.022	16.9354
COMPOUND WORDS	.879894	29.6585	129.502
UNIT MODIFIERS	11.8786	108.963	37.5036
WORD BREAKS	44.8746	207.074	18.9706
PAIRS OF PARENTHESES	2.63968	66.2906	103.242
PERIODS	58.0730	235.807	16.6932
TO TERMINATE	41.7950	200.164	19.6888
ABBREVIATIONS	16.2780	126.571	31.9659
DECIMAL POINTS	.439947	20.9749	196.000
QUESTION MARKS	0	0	0
TO TERMINATE	0	0	0
QUOTATION MARKS	19.3577	140.967	29.9378
OPEN QUOTATIONS	8.79894	93.4096	43.6433
CLOSE QUOTES	8.79894	93.4096	43.6433
SEMICOLONS	.439947	20.9749	196.000
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	0	0	0

TABLE D - 16

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM PAGE 1 OF THE CHICAGO TRIBUNE FOR FEBRUARY 27, 1966 (N = 2011)

MARK & USE	MEAN/1-WORD	STD.DEV.	+/- %
APOSTROPIES	17.9015	132.827	32.3610
POSSESSIVES	4.97265	70.3590	61.9417
CONTRACTIONS	12.9289	112.996	38.1990
COLONS	1.49180	38.6045	113.104
COMMAS	111.387	316.266	12.4099
SEPARATE QUOTES	1.98906	44.5656	97.9268
INTRODUCE PHRASES	4.47539	66.7051	65.2032
SEPARATE SERIES	8.45351	91.5703	47.3474
INDEPENDENT CLAUSES	8.95077	94.2076	46.0019
ADVERBIAL CLAUSES	3.97812	62.9624	69.1757
OPEN APPOSITIVES	1.49182	38.6045	113.104
CLOSE APPOSITIVES	1.49180	38.6045	113.104
OPEN PARENTHETICS	18.8961	136.192	31.5014
CLOSE PARENTHETICS	11.4371	106.357	40.6445
DASHES	2.48633	49.8134	87.5006
ELLIPSES	0	0	0
EXCLAMATION POINTS	1.49180	38.6045	113.104
TO TERMINATE	1.49100	38.6045	113.104
HYPHENS	34.8886	183.340	23.0209
COMPOUND WORDS	0	0	0
UNIT MODIFIERS	2.98359	54.5543	79.9171
WORD BREAKS	31.8250	175.578	24.1130
PAIRS OF PARENTHESES	0	0	0
PERIODS	58.1800	234.142	17.5896
TO TERMINATE	42.7648	202.377	20.6835
ABBREVIATIONS	13.9234	117.202	36.7909
DECIMAL POINTS	1.49180	38.6045	113.104
QUESTION MARKS	0	0	0
TO TERMINATE	0	0	0
QUOTATION MARKS	20.3879	141.358	30.3039
OPEN QUOTATIONS	8.95077	94.2076	46.0019
CLOSE QUOTES	7.45898	86.0641	50.4304
SEMICOLONS	0	0	0
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	0	0	0

TABLE D - 19

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM PAGE 1 OF THE ATLANTA JOURNAL AND CONSTITUTION FOR FEBRUARY 27, 1966 (N - 2378)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPIES	8.83095	93.5770	42.59ø4
POSSESSIVES	4.20521	64.7247	61.9632
CONTRACTIONS	4.62574	67.8695	58.9718
COLONS	.841043	28.9946	138.504
COMMAS	45.4163	208.259	18.4307
SEPARATE QUOTES	2.52313	50.1779	79.9325
INTRODUCE PHRASES	7.14886	84.2659	47.3707
SEPARATE SERIES	4.62574	67.8695	58.9718
INDEPENDENT CLAUSES	2.52313	50.1779	79.9325
ADVERBIAL CLAUSES	2.10261	45.8156	87.5801
OPEN APPOSITIVES	1.26156	35.5036	113.113
CLOSE APPOSITIVES	2.52313	50.1779	79.9325
OPEN PARENTHETICS	15.5593	123.789	31.9773
CLOSE PARENTHETICS	11.3541	105.971	37.5134
DASHES	2.10261	45.8156	87.5601
ELLIPSES	.420521	20.5066	196.000
EXCLAMATION POINTS	ø	ø	ø
TO TERMINATE	ø	ø	ø
HYPHENS	38.6880	192.891	20.3394
COMPOUND WORDS	2.10261	45.8156	87.5801
UNIT MODIFIERS	4.20521	64.7247	61.8632
WORD BREAKS	33.2212	179.251	21.6869
PAIRS OF PARENTHESES	ø	ø	ø
PERIODS	63.9193	244.660	15.3845
TO TERMINATE	18.3600	214.571	17.8334
ABBREVIATIONS	14.7183	120.448	32.8923
DECIMAL POINTS	.841ø43	28.9946	138.564
QUESTION MARKS	ø	ø	ø
TO TERMINATE	ø	ø	ø
QUOTATION MARKS	18.0824	136.398	30.3180
OPEN QUOTATIONS	9.67199	97.8901	40.6793
CLOSE QUOTES	7.56939	86.6905	45.0321
SEMICOLONS	ø	ø	ø
SEPARATE IND. CLAUSES	ø	ø	ø
SUPERIOR COMMAS	ø	ø	ø

TABLE D - 20

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM PAGE 1 OF THE DENVER POST FOR FEBRUARY 27, 1966 (N = 1681)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPIES	5.35396	72.9963	65.1770
POSSESSIVES	4.16419	64.4152	73.9480
CONTRACTIONS	1.18977	34.4827	138.552
COLONS	1.18977	34.4827	138.552
COMMAS	62.4628	244.513	18.7134
SEPARATE QUOTES	5.94884	76.9219	61.9144
INTRODUCE PHRASES	9.51814	97.1245	48.7808
SEPARATE SERIES	4.75907	68.9422	69.1519
INDEPENDENT CLAUSES	6.54372	80.6522	58.9201
ADVERBIAL CLAUSES	1.78465	42.2200	113.093
OPEN APPOSITIVES	5.35396	72.9963	65.1770
CLOSE APPOSITIVES	4.75907	68.9422	69.1519
OPEN PARENTHETICS	14.2772	118.667	39.7335
CLOSE PARENTHETICS	11.3028	105.744	44.7239
DASHES	.594884	24.3902	196.000
ELLIPSES	0	0	0
EXCLAMATION POINTS	0	0	0
TO TERMINATE	0	0	0
HYPHENS	43.4265	209.634	23.0772
COMPOUND WORDS	0	0	0
UNIT MODIFIERS	9.51814	97.1245	48.7805
WORD BREAKS	33.9084	181.747	25.5245
PAIRS OF PARENTHESES	2.97954	59.7140	119.965
PERIODS	65.4372	247.369	18.8715
TO TERMINATE	38.6675	192.859	23.9432
ABBREVIATIONS	27.3647	163.192	28.5000
DECIMAL POINTS	0	0	0
QUESTION MARKS	0	0	0
TO TERMINATE	0	0	0
QUOTATION MARKS	23.7954	152.456	30.8285
OPEN QUOTATIONS	10.1130	100.084	47.3101
CLOSE QUOTES	10.1130	100.084	47.3101
SEMICOLONS	.594884	24.3902	196.000
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	.594884	24.3902	196.000

TABLE D - 21

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM PAGE 1 OF THE NEW YORK TIMES FOR FEBRUARY 26, 1950 (N = 2910)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPIES	9.60854	97.5685	37.5452
POSSESSIVES	7.82918	88.1513	41.6209
CONTRACTIONS	.711744	26.6738	130.508
COLONS	2.49110	49.8576	74.0719
COMMAS	46.6102	212.541	16.8570
SEPARATE QUOTES	2.49110	49.8576	74.0719
INTRODUCE PHRASES	6.76157	81.9840	44.8212
SEPARATE SERIES	6.40560	79.7930	46.0576
INDEPENDENT CLAUSES	1.77936	42.1524	87.5914
ADVERBIAL CLAUSES	1.77936	42.1524	87.5914
OPEN APPOSITIVES	4.62633	67.8717	54.2444
CLOSE APPOSITIVES	4.27046	65.2207	56.4694
OPEN PARENTHETICS	13.5231	115.521	31.5853
CLOSE PARENTHETICS	9.60854	97.5685	37.5452
DASHES	3.55872	59.5594	61.8813
ELLIPSES	0	0	0
EXCLAMATION POINTS	0	0	0
TO TERMINATE	0	0	0
HYPHENS	51.6014	221.260	15.8542
COMPOUND WORDS	4.62633	67.8717	54.2444
UNIT MODIFIERS	8.18505	90.1163	40.7085
WORD BREAKS	38.7900	193.129	18.4090
PAIRS OF PARENTHESES	.355872	18.8646	196.000
PERIODS	68.6933	252.960	12.6177
TO TERMINATE	44.4810	206.204	17.1395
ABBREVIATIONS	24.1993	153.695	25.4833
DECIMAL POINTS	.711744	26.6738	130.508
QUESTION MARKS	.355872	18.8646	196.000
TO TERMINATE	.355872	18.8646	196.000
QUOTATION MARKS	28.4698	195.829	25.4329
OPEN QUOTATIONS	13.8790	117.018	31.1721
CLOSE QUOTES	12.4555	110.927	32.9289
SEMICOLONS	.355872	18.8646	196.000
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	.355872	18.8646	196.000

TABLE D - 22

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM
PAGE 1 OF THE LOS ANGELES TIMES
FOR FEBRUARY 26, 1950 (N = 2459)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPIES	8.94673	94.1822	41.6085
POSSESSIVES	6.91336	82.8757	47.3820
CONTRACTIONS	1.22001	34.9144	113.115
COLONS	.813330	28.5133	138.505
COMMAS	85.0671	251.593	15.2832
SEPARATE QUOTES	.813339	28.5133	138.505
INTRODUCE PHRASES	2.03335	45.0559	87.5825
SEPARATE SERIES	8.94673	94.1822	41.6085
INDEPENDENT CLAUSES	1.62668	40.3075	97.9402
ADVERBIAL CLAUSES	1.62668	40.3075	97.9402
OPEN APPOSITIVES	4.00003	63.7007	56.4526
CLOSE APPOSITIVES	4.47336	66.7470	58.9759
OPEN PARENTHETICS	20.3335	141.167	27.4409
CLOSE PARENTHETICS	13.4201	115.089	33.8964
DASHES	2.44002	49.3463	79.0352
ELLIPSES	0	0	0
EXCLAMATION POINTS	0	0	0
TO TERMINATE	0	0	0
HYPHENS	51.2403	220.532	17.9113
COMPOUND WORDS	.406669	20.1660	196.000
UNIT MODIFIERS	8.13339	89.8201	43.6572
WORD BREAKS	42.9203	204.959	18.4450
PAIRS OF PARENTHESES	1.22001	34.9144	113.115
PERIODS	70.3538	255.794	14.3709
TO TERMINATE	48.3937	214.640	17.5507
ABBREVIATIONS	21.7401	142.543	27.1969
DECIMAL POINTS	1.62668	40.3075	97.9402
QUESTION MARKS	0	0	0
TO TERMINATE	0	0	0
QUOTATION MARKS	8.94673	102.458	45.2646
OPEN QUOTATIONS	4.47336	66.7470	58.9759
CLOSE QUOTES	4.47336	66.7470	58.9759
SEMICOLONS	.406669	20.1660	196.000
SEPARATE IND. CLAUSES	.406669	20.1660	196.000
SUPERIOR COMMAS	0	0	0

TABLE D - 23

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM
PAGE 1 OF THE CHICAGO TRIBUNE
FOR FEBRUARY 26, 1956 (N = 2344)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPHES	9.39567	96.4445	41.5997
POSSESSIVES	5.97270	77.0685	52.2377
CONTRACTIONS	3.41297	58.3333	69.1929
COLONS	0	0	0
COMMAS	66.1263	251.967	15.4257
SEPARATE QUOTES	.853642	29.2041	136.503
INTRODUCE PHRASES	2.98635	54.5775	73.9801
SEPARATE SERIES	9.38567	96.4445	41.5997
INDEPENDENT CLAUSES	3.83959	61.8586	65.2217
ADVERBIAL CLAUSES	1.27986	35.7599	113.112
OPEN APPOSITIVES	6.82594	82.3543	48.8429
CLOSE APPOSITIVES	5.11945	71.3822	56.4474
OPEN PARENTHETICS	23.0375	152.055	26.3688
CLOSE PARENTHETICS	14.9317	121.300	32.9898
DASHES	2.13311	46.1461	87.5790
ELLIPSES	0	0	0
EXCLAMATION POINTS	.853642	29.2041	136.503
TO TERMINATE	.426821	20.6548	196.000
HYPHENS	37.9693	193.383	20.6187
COMPOUND WORDS	1.27986	35.7599	113.112
UNIT MODIFIERS	4.26621	65.1967	61.8615
WORD BREAKS	32.4232	177.159	23.1200
PAIRS OF PARENTHESES	2.55973	50.5397	79.3316
PERIODS	71.2457	257.290	14.9198
TO TERMINATE	48.2082	214.252	17.9920
ABBREVIATIONS	23.4642	151.405	26.1423
DECIMAL POINTS	.426821	20.6548	196.000
QUESTION MARKS	0	0	0
TO TERMINATE	0	0	0
QUOTATION MARKS	14.5051	119.586	33.3762
OPEN QUOTATIONS	5.11945	71.3822	56.4474
CLOSE QUOTES	5.11945	71.3822	56.4474
SEMICOLONS	1.27986	35.7599	113.112
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	.853642	29.2041	136.503

TABLE D - 24

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM PAGE 1 OF THE ATLANTA JOURNAL AND CONSTITUTION FOR FEBRUARY 26, 1956 (N = 3132)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPIES	11.8135	108.0653	36.0364
POSSESSIVES	8.301400	96.7476	38.2850
CONTRACTIONS	2.871356	53.5371	65.2498
COLONS	.957854	30.9393	113.125
COMMAS	59.0677	237.140	14.06835
SEPARATE QUOTES	2.554288	50.4833	69.2190
INTRODUCE PHRASES	3.831427	61.7896	50.4809
SEPARATE SERIES	9.997831	99.01101	35.0330
INDEPENDENT CLAUSES	1.596427	39.9297	87.5979
ADVERBIAL CLAUSES	1.915711	43.7336	79.9527
OPEN APPOSITIVES	3.831427	61.7896	50.4808
CLOSE APPOSITIVES	3.192855	56.4247	61.8915
OPEN PARENTHETICS	17.56007	131.369	26.1997
CLOSE PARENTHETICS	11.17500	105.136	36.9497
DASHES	2.234994	47.2304	74.0100
ELLIPSES	.638570	25.2659	138.571
EXCLAMATION POINTS	0	0	0
TO TERMINATE	0	0	0
HYPHENS	41.50700	199.492	16.83425
COMPOUND WORDS	1.277140	35.7200	97.9530
UNIT MODIFIERS	5.108556	71.3029	48.9825
WORD BREAKS	34.16350	181.678	18.62415
PAIRS OF PARENTHESES	2.554288	50.4833	69.2190
PERIODS	70.24277	265.596	12.74380
TO TERMINATE	51.08556	220.207	15.09966
ABBREVIATIONS	20.11499	140.416	24.44880
DECIMAL POINTS	0	0	0
QUESTION MARKS	.957854	30.9393	113.125
TO TERMINATE	.957854	30.9393	113.125
QUOTATION MARKS	16.92221	133.860	27.70440
OPEN QUOTATIONS	7.662844	87.2155	39.8611
CLOSE QUOTES	7.343555	85.3929	40.7250
SEMICOLONS	.319285	17.8685	196.000
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	0	0	0

TABLE D - 25

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM
PAGE 1 OF THE DENVER POST
FOR FEBRUARY 26, 1956 (N = 2445)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPIES	8.17996	90.09909	43.65562
POSSESSIVES	4.90796	69.89991	56.45529
CONTRACTIONS	3.27198	57.11993	69.19172
COLONS	.817996	28.5948	138.5065
COMMAS	49.8978	225.168	17.88172
SEPARATE QUOTES	0	0	0
INTRODUCE PHRASES	2.04499	45.1845	87.56221
SEPARATE SERIES	7.30196	85.5630	46.03367
INDEPENDENT CLAUSES	1.63599	40.4226	97.93998
ADVERBIAL CLAUSES	1.63599	40.4226	97.93998
OPEN APPOSITIVES	2.45399	49.4879	79.93348
CLOSE APPOSITIVES	2.04499	45.1845	87.58221
OPEN PARENTHETICS	16.7680	129.431	30.35065
CLOSE PARENTHETICS	8.99796	94.4491	41.56074
DASHES	4.49898	66.9371	58.97552
ELLIPSES	0	0	0
EXCLAMATION POINTS	0	0	0
TO TERMINATE	0	0	0
HYPHENS	29.4479	169.093	22.76008
COMPOUND WORDS	.817996	28.5948	138.5065
UNIT MODIFIERS	9.40695	96.5519	40.69465
WORD BREAKS	20.0400	140.169	27.72737
PAIRS OF PARENTHESES	1.63599	40.4226	97.93998
PERIODS	52.3517	222.781	16.86980
TO TERMINATE	40.0816	196.191	19.40221
ABBREVIATIONS	11.9800	108.292	36.18772
DECIMAL POINTS	.408998	20.2237	196
QUESTION MARKS	.408998	20.2237	196
TO TERMINATE	0	0	0
QUOTATION MARKS	1.63599	40.5036	110.3041
OPEN QUOTATIONS	.817996	28.5948	138.5065
CLOSE QUOTES	.817996	28.5948	138.5065
SEMICOLONS	.817996	28.5948	138.5065
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	0	0	0

TABLE D - 28

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM
PAGE 1 OF THE NEW YORK TIMES
FOR MARCH 3, 1946 (N = 3949)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPIES	8.35655	91.0429	33.9807
POSSESSIVES	7.59686	86.8393	35.6529
CONTRACTIONS	0	0	0
COLONS	4.05166	63.5316	48.9069
COMMAS	52.6716	237.688	14.0748
SEPARATE QUOTES	.253229	15.9132	196.000
INTRODUCE PHRASES	3.29197	57.2985	54.2779
SEPARATE SERIES	5.57103	74.4406	41.6761
INDEPENDENT CLAUSES	3.79843	61.5220	50.5172
ADVERBIAL CLAUSES	1.77260	42.0703	74.0247
OPEN APPOSITIVES	4.05166	63.5316	48.9069
CLOSE APPOSITIVES	2.78552	52.7111	59.0213
OPEN PARENTHETICS	14.6873	120.313	25.5496
CLOSE PARENTHETICS	8.35655	91.0429	33.9807
DASHES	.759686	27.5554	113.132
ELLIPSES	0	0	0
EXCLAMATION POINTS	0	0	0
TO TERMINATE	0	0	0
HYPHENS	47.6070	214.146	14.0298
COMPOUND WORDS	1.51937	38.9544	79.9660
UNIT MODIFIERS	5.06457	70.9944	43.7213
WORD BREAKS	41.0230	198.369	15.0820
PAIRS OF PARENTHESES	3.79843	61.5220	50.5172
PERIODS	52.1651	222.369	13.2967
TO TERMINATE	37.7211	190.569	15.7531
ABBREVIATIONS	10.9888	103.793	29.7303
DECIMAL POINTS	0	0	0
QUESTION MARKS	0	0	0
TO TERMINATE	0	0	0
QUOTATION MARKS	13.9276	127.554	28.5043
OPEN QUOTATIONS	5.82428	76.1039	40.7548
CLOSE QUOTES	5.57103	74.4406	41.6761
SEMICOLONS	.253229	15.9132	196.000
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	0	0	0

TABLE D - 27

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM
PAGE 1 OF THE LOS ANGELES TIMES
FOR MARCH 3, 1946 (N = 3210)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPIES	6.85358	82.5150	41.6504
POSSESSIVES	5.23053	78.6997	43.6979
CONTRACTIONS	.623053	24.9571	139.571
COLONS	.934579	30.5614	113.125
COMMAS	49.8442	217.857	15.1064
SEPARATE QUOTES	2.49221	49.8676	69.2208
INTRODUCE PHRASES	4.67290	68.2093	50.4065
SEPARATE SERIES	3.73832	61.0369	56.4833
INDEPENDENT CLAUSES	1.55763	39.4422	87.5992
ADVERBIAL CLAUSES	.934579	30.5614	113.125
OPEN APPOSITIVES	3.42679	58.4475	59.3041
CLOSE APPOSITIVES	1.55763	39.4422	87.5992
OPEN PARENTHETICS	18.6916	135.455	25.3690
CLOSE PARENTHETICS	13.3956	114.998	29.6935
DASHES	2.80374	52.8843	65.2518
ELLIPSES	0	0	0
EXCLAMATION POINTS	0	0	0
TO TERMINATE	0	0	0
HYPHENS	65.1000	248.016	13.1778
COMPOUND WORDS	1.86916	43.2001	79.9543
UNIT MODIFIERS	11.2150	105.322	32.4880
WORD BREAKS	50.4673	218.941	15.0079
PAIRS OF PARENTHESES	1.24611	35.2837	97.9542
PERIODS	56.5670	234.849	12.8720
TO TERMINATE	37.3832	189.728	17.5574
ABBREVIATIONS	21.1838	144.019	23.5191
DECIMAL POINTS	0	0	0
QUESTION MARKS	0	0	0
TO TERMINATE	0	0	0
QUOTATION MARKS	15.8879	132.326	28.8126
OPEN QUOTATIONS	7.78816	87.9199	39.8531
CLOSE QUOTES	7.47664	86.1571	39.8647
SEMICOLONS	.311526	17.6501	196.000
SEPARATE IND. CLAUSES	0	0	0
SUPERIOR COMMAS	.311526	17.6501	196.000

TABLE D - 28

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM PAGE 1 OF THE CHICAGO TRIBUNE FOR MARCH 3, 1946 (N = 3014)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPIES	11.3448	105.921	30.4401
POSSESSIVES	10.2380	100.677	32.0613
CONTRACTIONS	.553403	23.5213	136.574
COLONS	1.38351	37.1749	87.6053
COMMAS	60.3210	240.427	12.9950
SEPARATE QUOTES	1.93691	43.9738	74.0195
INTRODUCE PHRASES	4.98063	70.4074	46.0088
SEPARATE SERIES	10.7914	103.334	31.2196
INDEPENDENT CLAUSES	2.49032	49.8478	65.2610
ADVERBIAL CLAUSES	1.66021	40.7175	79.9613
OPEN APPOSITIVES	4.42723	66.3991	48.8982
CLOSE APPOSITIVES	4.42723	66.3991	48.8982
OPEN PARENTHETICS	16.8788	128.935	24.8960
CLOSE PARENTHETICS	11.3448	105.921	30.4401
DASHES	1.93691	43.9738	74.0195
ELLIPSES	0	0	0
EXCLAMATION POINTS	0	0	0
TO TERMINATE	0	0	0
HYPHENS	33.4809	179.914	17.5198
COMPOUND WORDS	.830105	28.8036	113.129
UNIT MODIFIERS	1.10691	33.2549	97.9593
WORD BREAKS	31.2673	174.863	18.1501
PAIRS OF PARENTHESES	1.10691	33.2549	97.9593
PERIODS	54.7869	227.595	13.5440
TO TERMINATE	40.1217	196.272	15.9402
ABBREVIATIONS	15.4953	123.529	25.9915
DECIMAL POINTS	.830105	28.8036	113.129
QUESTION MARKS	0	0	0
TO TERMINATE	0	0	0
QUOTATION MARKS	20.7526	160.820	25.2650
OPEN QUOTATIONS	8.57775	92.7308	35.4562
CLOSE QUOTES	7.47095	86.1251	37.5843
SEMICOLONS	.553403	23.5213	136.574
SEPARATE IND. CLAUSES	.276702	16.6344	196.000
SUPERIOR COMMAS	.276702	16.6344	196.000

TABLE D - 29

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM
PAGE 1 OF THE ATLANTA CONSTITUTION
FOR MARCH 3, 1946 (N = 3131)

MARK & USE	MEAN/M-WORD	STD. DEV.	+/- %
APOSTROPIES	13.0949	113.699	30.4136
POSSESSIVES	8.340406	90.7619	38.2849
CONTRACTIONS	3.883264	61.7994	50.4808
COLONS	.319387	17.8714	196.000
COMMAS	57.8090	233.419	14.1435
SEPARATE QUOTES	2.87448	53.5456	65.2498
INTRODUCE PHRASES	2.23571	47.2379	74.0100
SEPARATE SERIES	5.74896	75.8157	46.0720
INDEPENDENT CLAUSES	2.55509	50.4914	69.2109
ADVERBIAL CLAUSES	3.51365	59.1780	59.0017
OPEN APPOSITIVES	5.11019	71.3141	48.8824
CLOSE APPOSITIVES	3.883264	61.7994	56.4808
OPEN PARENTHETICS	19.9020	139.242	24.6483
CLOSE PARENTHETICS	13.4142	115.859	30.8447
DASHES	1.916326	43.7408	79.9527
ELLIPSES	.638774	25.2699	138.571
EXCLAMATION POINTS	0	0	0
TO TERMINATE	0	0	0
HYPHENS	40.8815	198.047	16.9690
COMPOUND WORDS	.958160	30.9443	113.124
UNIT MODIFIERS	9.58160	97.4311	35.6184
WORD BREAKS	30.6611	172.425	19.6902
PAIRS OF PARENTHESES	.638774	25.2699	138.571
PERIODS	55.8927	229.751	14.3985
TO TERMINATE	40.8815	198.047	16.9690
ABBREVIATIONS	14.0530	117.729	29.3444
DECIMAL POINTS	.958160	30.9443	113.124
QUESTION MARKS	0	0	0
TO TERMINATE	0	0	0
QUOTATION MARKS	22.3571	150.301	24.8716
OPEN QUOTATIONS	10.2204	100.594	34.4762
CLOSE QUOTES	9.58160	97.4311	35.6184
SEMICOLONS	1.27755	35.7257	97.9530
SEPARATE IND. CLAUSES	1.27755	35.7257	97.9530
SUPERIOR COMMAS	0	0	0

TABLE D - 30

PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FOUR
PAGE 1 OF THE DENVER POST
FOR MARCH 3, 1946 (N = 3047)

MARK & USE	MEAN/K-WORD	STD.DEV.	+/- %
APOSTROPHIES	8.311716	90.83110	35.66404
POSSESSIVES	6.337649	79.66020	40.74400
CONTRACTIONS	1.966343	40.7569	79.96126
COLONS	2.321791	47.04300	69.22092
COMMAS	70.14114	256.5014	11.99345
SEPARATE QUOTES	1.338619	37.32110	87.66052
INTRODUCE PHRASES	4.713006	68.40992	47.43114
SEPARATE SERIES	19.40067	137.9069	23.32013
INDEPENDENT CLAUSES	4.713006	68.40992	47.43114
ADVERBIAL CLAUSES	.5544477	23.55441	138.5741
OPEN APPOSITIVES	5.544477	74.32668	43.71113
CLOSE APPOSITIVES	4.713006	68.40992	47.43114
OPEN PARENTHETICS	15.80026	124.7268	25.75065
CLOSE PARENTHETICS	11.99213	108.5477	29.71511
DASHES	4.713006	68.40992	47.43114
ELLIPSES	0	0	0
EXCLAMATION POINTS	.2772739	16.66505	196.0000
TO TERMINATE	0	0	0
HYPHENS	39.09907	193.8338	16.18126
COMPOUND WORDS	2.321791	47.04300	69.22924
UNIT MODIFIERS	12.19185	100.7066	29.37114
WORD BREAKS	24.66742	155.1552	20.52000
PAIRS OF PARENTHESES	1.940067	44.01164	74.01194
PERIODS	52.10115	241.3273	12.66844
TO TERMINATE	37.99817	191.1794	16.42666
ABBREVIATIONS	22.73336	149.0074	21.40001
DECIMAL POINTS	1.338619	37.32110	87.66052
QUESTION MARKS	.5544477	23.55441	138.5741
TO TERMINATE	.2772739	16.66505	196.0000
QUOTATION MARKS	16.66343	127.9115	25.00956
OPEN QUOTATIONS	6.993097	82.99749	39.06933
CLOSE QUOTES	7.485544	86.20660	37.50440
SEMICOLONS	1.940067	44.01164	74.01194
SEPARATE IND. CLAUSES	1.108895	33.29271	97.99592
SUPERIOR COMMAS	.8311716	28.83115	113.1129

APPENDIX E

STATISTICAL ABSTRACTS AND COMPUTER PROGRAM

```
10 REMARKS: PGRM TO CALCULATE PUNCT. FREQ., STD.DEV. & CONF. LIMITS
15 REM: WITHOUT ECHO CHECK OF DATA IN PRINTOUT
20 REM: AS A % TOLLERANCE FROM SUM OF X'S & SUM OF X'S SQD.
30 REM: AND PRINT 40 TABLES.
40 REM: NAMES OF VARABLES FOLLOW:
50 REM: C1=TBL NOS.: C2=LINE NOS.
70 REM: M1=MEAN FREQ./WORD; M2=MEAN FREQ./KILOWORD
80 REM: N1=SAMPLE SIZE IN CELL; N2=NEWSPAPER PAGE NUMBER
R+120 REM: T1=ABS. CONF. LIMITS(+/-): T2=CONF.LIMITS AS % OF MEAN
1000 REM: CONTINUE
1020 FOR C1=1 TO 30 STEP 1
1040 READ N1
1050 PRINT " "
1051 PRINT " "
1052 PRINT " "
1053 PRINT " "
1054 PRINT " "
1055 PRINT " "
1056 PRINT " "
1064 PRINT " "
1065 PRINT " "
1066 PRINT " "
1067 PRINT " "
1068 PRINT TAB(35)"TABLE D -"C1
1069 PRINT " "
1090 PRINT TAB(10)"PUNCTUATION FREQUENCY BY MARK AND USE OF MARK FROM
1100 IF C1<11 GO TO 1200
1110 IF C1>15 GO TO 1200
1120 LET N2=2
1130 GO TO 1210
1200 LET N2=1
1210 REM: CONTINUE
1220 LET K1=C1*0.2 + 0.800000001
1230 LET K2=INT(K1)
1240 LET K3=K1-K2
1250 IF K3<0.1 GO TO 1400
1260 IF K3<0.3 GO TO 1420
1270 IF K3<0.5 GO TO 1440
1280 IF K3<0.7 GO TO 1460
1290 IF K3<0.9 GO TO 1480
1400 PRINT TAB(10)"PAGE "N2" OF THE NEW YORK TIMES"
1410 GO TO 1500
1420 PRINT TAB(10)"PAGE "N2" OF THE LOS ANGELES TIMES"
1430 GO TO 1500
1440 PRINT TAB(10)"PAGE "N2" OF THE CHICAGO TRIBUNE"
1450 GO TO 1500
1460 IF C1>25 GO TO 1465
1461 PRINT TAB(10)"PAGE "N2" OF THE ATLANTA JOURNAL AND CONSTITUTION"
1462 GO TO 1500
1465 PRINT TAB(10)"PAGE "N2" OF THE ATLANTA CONSTITUTION"
1466 GO TO 1500
```

```
1480 PRINT TAB(10)"PAGE "N2" OF THE DENVER POST"
1490 GO TO 1500
1500 REM: CONTINUE
1510 IF K2<4 GO TO 1600
1520 IF K2=4 GO TO 1620
1530 IF K2=5 GO TO 1640
1540 IF K2=6 GO TO 1660
1600 PRINT TAB(10)"FOR FEBRUARY 29, 1976 (N = "N1")"
1610 GO TO 1700
1620 PRINT TAB(10)"FOR FEBRUARY 27, 1966 (N = "N1")"
1630 GO TO 1700
1640 PRINT TAB(10)"FOR FEBRUARY 26, 1956 (N = "N1")"
1650 GO TO 1700
1660 PRINT TAB(10)"FOR MARCH 3, 1946 (N = "N1")"
1700 REM: CONTINUE
1802 PRINT " "
1810 PRINT "MARK #","USE","MEAN/K-WORD","STD.DEV.","+/- %"
1820 PRINT " "
1830 FOR C2=1 TO 35 STEP 1
1840 READ X1,X2
1841 IF X1=0 GO TO 1920
1842 IF X2=0 GO TO 1920
1850 LET M1=X1/N1
1860 LET M2=1000*M1
1870 LET S1=((X2-N1*M1**2)/(N1-1))**0.5
1880 LET S2=1000*S1
1890 LET T1=1.96*S1/N1**0.5
1895 LET T2=100*T1/M1
1900 GO TO 2000
1920 LET M2=0
1940 LET S2=0
1960 LET T2=0
2000 IF C2=1 GO TO 2000
2020 IF C2=2 GO TO 3020
2040 IF C2=3 GO TO 3040
2060 IF C2=4 GO TO 3060
2080 IF C2=5 GO TO 3080
2100 IF C2=6 GO TO 3100
2120 IF C2=7 GO TO 3120
2140 IF C2=8 GO TO 3140
2160 IF C2=9 GO TO 3160
2180 IF C2=10 GO TO 3180
2200 IF C2=11 GO TO 3200
2220 IF C2=12 GO TO 3220
2240 IF C2=13 GO TO 3240
2260 IF C2=14 GO TO 3260
2280 IF C2=15 GO TO 3280
2300 IF C2=16 GO TO 3300
2320 IF C2=17 GO TO 3320
2340 IF C2=18 GO TO 3340
2360 IF C2=19 GO TO 3360
```

```
2380 IF C2=20 GO TO 3380
2400 IF C2=21 GO TO 3400
2440 IF C2=22 GO TO 3440
2460 IF C2=23 GO TO 3460
2480 IF C2=24 GO TO 3480
2500 IF C2=25 GO TO 3500
2520 IF C2=26 GO TO 3520
2540 IF C2=27 GO TO 3540
2560 IF C2=28 GO TO 3560
2580 IF C2=29 GO TO 3580
2600 IF C2=30 GO TO 3600
2620 IF C2=31 GO TO 3620
2640 IF C2=32 GO TO 3640
2660 IF C2=33 GO TO 3660
2680 IF C2=34 GO TO 3680
2690 IF C2=35 GO TO 3700
3000 PRINT "APOSTROPIES",  ",M2,S2,T2
3010 GO TO  4000
3020 PRINT "  POSSESSIVES",M2,S2,T2
3030 GO TO  4000
3040 PRINT "  CONTRACTIONS",M2,S2,T2
3041 PRINT " "
3050 GO TO  4000
3060 PRINT "COLONS",  ",M2,S2,T2
3061 PRINT " "
3070 GO TO  4000
3080 PRINT "SEMMAS",  ",M2,S2,T2
3090 GO TO  4000
3100 PRINT "  SEPARATE QUOTES",M2,S2,T2
3110 GO TO  4000
3120 PRINT "  INTRODUCE PHRASES",M2,S2,T2
3130 GO TO  4000
3140 PRINT "  SEPARATE SERIES",M2,S2,T2
3150 GO TO  4000
3160 PRINT "  INDEPENDENT CLAUSES",M2,S2,T2
3170 GO TO  4000
3180 PRINT "  ADVERBIAL CLAUSES",M2,S2,T2
3190 GO TO  4000
3200 PRINT "  OPEN APPOSITIVES",M2,S2,T2
3210 GO TO  4000
3220 PRINT "  CLOSE APPOSITIVES",M2,S2,T2
3230 GO TO  4000
3240 PRINT "  OPEN PARENTHETICS",M2,S2,T2
3250 GO TO  4000
3260 PRINT "  CLOSE PARENTHETICS",M2,S2,T2
3261 PRINT " "
3270 GO TO  4000
3280 PRINT "DASHES",  ",M2,S2,T2
3281 PRINT " "
3290 GO TO  4000
3300 PRINT "ELLIPSES",  ",M2,S2,T2
3301 PRINT " "
```

```
3310 GO TO 4000
3320 PRINT "EXCLAMATION POINTS",M2,S2,T2
3330 GO TO 4000
3340 PRINT "  TO TERMINATE",M2,S2,T2
3341 PRINT " "
3350 GO TO 4000
3360 PRINT "HYPHENS"," ",M2,S2,T2
3370 GO TO 4000
3380 PRINT "  COMPOUND WORDS",M2,S2,T2
3390 GO TO 4000
3400 PRINT "  UNIT MODIFIERS",M2,S2,T2
3410 GO TO 4000
3440 PRINT "  WORD BREAKS",M2,S2,T2
3441 PRINT " "
3450 GO TO 4000
3460 PRINT "PAIRS OF PARENTHESES",M2,S2,T2
3461 PRINT " "
3470 GO TO 4000
3480 PRINT "PERIODS"," ",M2,S2,T2
3490 GO TO 4000
3500 PRINT "  TO TERMINATE",M2,S2,T2
3510 GO TO 4000
3520 PRINT "  ABBREVIATIONS",M2,S2,T2
3530 GO TO 4000
3540 PRINT "  DECIMAL POINTS",M2,S2,T2
3541 PRINT " "
3550 GO TO 4000
3560 PRINT "QUESTION MARKS",M2,S2,T2
3570 GO TO 4000
3580 PRINT "  TO TERMINATE",M2,S2,T2
3581 PRINT " "
3590 GO TO 4000
3600 PRINT "QUOTATION MARKS",M2,S2,T2
3610 GO TO 4000
3620 PRINT "  OPEN QUOTATIONS",M2,S2,T2
3630 GO TO 4000
3640 PRINT "  CLOSE QUOTES",M2,S2,T2
3641 PRINT " "
3650 GO TO 4000
3660 PRINT "SEMICOLONS"," ",M2,S2,T2
3670 GO TO 4000
3680 PRINT "  SEPARATE IND. CLAUSES",M2,S2,T2
3690 GO TO 4000
3700 PRINT "  SUPERIOR COMMAS",M2,S2,T2
3701 PRINT " "
3710 GO TO 4000
4000 REM :CONTINUE
4010 NEXT C2
4020 NEXT C1
4030 STOP
4040 END
```

```
5000 REM: NIT300P1,1976
5010 DATA 300
5020 DATA 1,1,1,1,0,0
5030 DATA 0,0
5040 DATA 16,16,0,0,0,0,2,2,0,0,0,0,4,4,0,0,10,10,0,0
5050 DATA 3,3
5060 DATA 0,0
5070 DATA 0,0,0,0
5080 DATA 13, 13,1,1,0,0,13,13
5090 DATA 0,0
5095 DATA 14,14,7,7,5,5,1,1
5100 DATA 0,0,0,0
5110 DATA 4,4,0,0,0,0
5120 DATA 1,1,0,0,1,1
5200 REM IAT300P1,1976
5210 DATA 300
5220 DATA 5,5,5,5,0,0
5230 DATA 0,0
5240 DATA 0,0,1,1,1,1,0,0,0,0,1,1,2,2,0,0,4,4,0,0
5250 DATA 0,0
5260 DATA 0,0
5270 DATA 0,0,0,0
5280 DATA 17,17,0,0,4,4,13,13
5290 DATA 1,1
5300 DATA 20,20,12,12,8,8,0,0
5310 DATA 0,0,0,0
5320 DATA 10,10,6,6,4,4
5330 DATA 0,0,0,0,0,0
5400 REM: CHI300P1,1976
5410 DATA 300
5420 DATA 2,2,2,2,0,0
5430 DATA  0,0
5440 DATA 10,10,3,3,2,2,0,0,0,0,0,0,2,2,0,0,2,2,0,0
5450 DATA 1,1
5460 DATA 0,0
5470 DATA 0,0,0,0
5480 DATA 14,14,10,10,0,0,4,4
5490 DATA 0,0
5500 DATA 12,12,10,10,2,2,0,0
5510 DATA 0,0,0,0
5520 DATA 7,7,3,3,3,3
5530 DATA 0,0,0,0,0,0
5600 REM:ATL300P1,1976
5605 DATA 300
5610 DATA 3,3,1,1,2,2
5620 DATA 0,0
5630 DATA 16,16,1,1,5,5,0,0,3,3,0,0,1,1,0,0,4,4,0,0
5640 DATA 0,0
5650 DATA 0,0
5660 DATA 0,0,0,0
5670 DATA 18,18,0,0,1,1,16,16
```

```
5680  DATA 0,0
5690 DATA  14,14,13,13,1,1,0,0
5700 DATA  0,0,0,0
5710 DATA  2,2,1,1,1,1
5720 DATA  0,0,0,0,0,0
5800 REM:DEN300P1,1976
5801 DATA  300
5802 DATA  6,6,5,5,1,1
5803 DATA  0,0
5804 DATA  12,12,0,0,1,1,4,4,1,1,0,0,0,0,0,0,4,4,0,0
5805 DATA  2,2
5806 DATA  0,0
5807 DATA  0,0,0,0
5808 DATA  11,11,1,1,0,0,10,10
5809 DATA  1,1
5810 DATA  18,18,11,11,0,0,7,7
5811 DATA  0,0,0,0
5812 DATA  0,0,0,0,0,0
5820 DATA  0,0,0,0,0,0
5900 REM:NYTALLP1,1976
5905 DATA  2135
5910 DATA  19,19,19,19,0,0
5920 DATA  3,3
5930 DATA  121,121,4,4,14,14,14,14,2,2,4,4,14,14,1,1,33,33,20,20
5940 DATA  9,9
5950 DATA  0,0
5960 DATA  1,1,0,0
5970 DATA  113,113,3,3,13,13,92,92
5980 DATA  0,0
5990 DATA  114,114,74,74,38,38,1,1
6000 DATA  0,0,0,0
6010 DATA  36,48,17,17,16,16
6020 DATA  0,0,0,0,0,0
6100 REM:LATALLP1,1976
6105 DATA  2201
6110 DATA  22,22,19,19,4,4
6120 DATA  3,3
6130 DATA  94,94,10,10,11,11,8,8,4,4,6,6,9,9,0,0,48,48,0,0
6140 DATA  10,10
6150 DATA  0,0
6160 DATA  0,0,0,0
6170 DATA  114,126,10,10,17,17,85,85
6180 DATA  6,10
6190 DATA  114,114,78,78,21,21,15,15
6200 DATA  0,0,0,0
6210 DATA  36,38,18,18,15,15
6220 DATA  3,3,0,0,0,0
6300 REM:CHIALLP1,1976
6305 DATA  2107
6310 DATA  33,33,24,24,4,4
6315 DATA  2,2
```

```
6320 DATA 116,116,13,13,13,13,10,10,3,3,4,4,12,12,6,6,38,38,20,20
6330 DATA 5,5
6340 DATA 0,0
6350 DATA 1,1,1,1
6360 DATA 92,92,2,2,10,10,86,86
6370 DATA 0,0
6380 DATA 134,134, 102,102,40,40,2,2
6390 DATA 0,0,0,0
6400 DATA 54,56,21,21,20,20
6410 DATA 3,3,0,0,0,0
6500 REM:ATLALLP1,1976
6505 DATA 1640
6510 DATA 19,19,9,9,13,13
6520 DATA 1,1
6530 DATA 82,82,3,3,17,17,16,16,17,17,3,3,4,4,3,3,17,17,6,6
6540 DATA 4,4
6550 DATA 0,0
6560 DATA 0,0,0,0
6570 DATA 70,70,0,0,12,12,57,57
6580 DATA 2,2
6590 DATA 103,103, 87,87,14,14,5,5
6600 DATA 1,1,1,1
6610 DATA 21,25, 6,6,5,5
6620 DATA 0,0,0,0,0,0
6700 REM: DENALLP1,1976
6705 DATA 2671
6710 DATA 40,40,20,20,17,17
6720 DATA 6,6
6730 DATA 143,143,9,9,25,25,22,22,5,5,3,3,9,9,5,5,38,38,21,21
6740 DATA 10,10
6750 DATA 0,0
6760 DATA 0,0,0,0
6770 DATA 98,98,4,4,19,19,72,72
6780 DATA 5,5
6790 DATA 146,146,105,105,30,30,11,11
6791 DATA 3,3,3,3
6792 DATA 65,77,26,26,25,25
6793 DATA 0,0,0,0,0,0
6800 REM: NYTALLP2,1976
6805 DATA 797
6810 DATA 9,9,9,9,0,0
6820 DATA 2,2
6830 DATA 37,37,2,2,0,0,0,0,1,1,2,2,11,11,5,5,6,6,3,3
6840 DATA 2,2
6850 DATA 0,0
6860 DATA 0,0,0,0
6870 DATA 32,34,1,1,4,4,22,22
6880 DATA 0,0
6890 DATA 57,57,38,38,16,16,2,2
6900 DATA 0,0,0,0
6910 DATA 21,34,5,5,6,6
```

```
6920 DATA 0,0,0,0,0,0
7000 REM:LATP2,1976
7005 DATA 923
7010 DATA 7,7,7,7,0,0
7020 DATA 0,0
7030 DATA 39,39,1,1,1,1,3,3,4,4,1,1,1,1,1,1,16,16,2,8
7040 DATA 0,0
7050 DATA 1,1
7060 DATA 0,0,0,0
7070 DATA 54,54,4,4,3,3,48,48
7080 DATA 0,0
7090 DATA 48,48,36,36,12,12,0,0
7100 DATA 0,0,0,0
7110 DATA 8,8,4,4,4,4
7120 DATA 0,0,0,0,0,0
7200 REM: CHIP2,1976
7205 DATA 966
7210 DATA 9,9,5,5,4,4
7220 DATA 0,0
7230 DATA 75,75,10,10,10,10,0,0,5,5,0,0,4,4,3,3,30,30,15,15
7240 DATA 2,2
7250 DATA 0,0
7260 DATA 1,1,0,0
7270 DATA 34,34,1,1,4,4,31,31
7280 DATA 0,0
7290 DATA 60,60,41,41,19,19,1,1
7300 DATA 0,0,0,0
7310 DATA 27,35,14,14,11,11
7320 DATA 0,0,0,0,0,0
7400 REM: ATLP2,1976
7405 DATA 1353
7410 DATA 18,18,16,16,2,2
7420 DATA 2,2
7430 DATA 90,92,7,7,16,16,18,18,6,6,3,3,9,9,5,5,14,14,4,4
7440 DATA 4,4
7450 DATA 0,0
7460 DATA 0,0,0,0
7470 DATA 52,52,0,0,1,1,52,52
7480 DATA 1,1
7490 DATA 75,75,54,54,18,18,4,4
7500 DATA 0,0,0,0
7510 DATA 34,34,16,16,14,14
7520 DATA 0,0,0,0,0,0
7600 REM:DENP2,1976
7605 DATA 341
7610 DATA 6,6,4,4,2,2
7620 DATA 4,4
7630 DATA 42,42,1,1,1,1,26,26,3,3,3,3,0,0,7,7,5,5,2,2
7640 DATA 0,0
7650 DATA 0,0
7660 DATA 0,0,0,0
```

```
7670 DATA 11,13,1,1,3,3,5,5
7680 DATA 0,0
7690 DATA 19,19,16,16,1,1,0,0
7700 DATA 1,1,1,1
7710 DATA 6,6,2,2,2,2
7720 DATA 0,0,0,0,0,0
7800 REM:NYTP1,1966
7805 DATA 2377
7810 DATA 24,24,21,21,3,3
7820 DATA 2,2
7830 DATA 117,117,7,7,9,9,12,12,4,4,4,4,16,16,8,8,33,33,17,17
7840 DATA 0,0
7850 DATA 0,0
7860 DATA 0,0,0,0
7870 DATA 139,139,0,0,17,17,120,120
7880 DATA 0,0
7890 DATA 143,143,94,94,47,47,1,1
7900 DATA 0,0,0,0
7910 DATA 60,60,29,29,27,27
7920 DATA 3,3,0,0,3,3
8000 REM: LATP1,1966
8005DATA 2273
8010 DATA 15,15,15,15,0,0
8020 DATA 13,13
8030 DATA 95,95,1,1,12,12,0,0,4,4,4,4,0,0,0,0,29,29,15,15
8040 DATA 8,8
8050 DATA 0,0
8060 DATA 0,0,0,0
8070 DATA 130,134,2,2,27,27,102,102
8085 DATA 6,10
8090 DATA 132,134,95,95,37,37,1,1
8091 DATA 0,0,0,0
8092 DATA 44,46,20,20,20,20
8093 DATA 1,1,0,0,0,0
8200 REM:CHIP1,1966
8105 DATA 2011
8110 DATA 36,36,10,10,26,26
8115 DATA 3,3
8120 DATA 224,226,4,4,9,9,17,17,18,18,8,8,3,3,3,3,38,38,23,23
8125 DATA 5,5
8130 DATA 0,0
8135 DATA 3,3,3,3
8140 DATA 70,70,0,0,6,6,64,64
8145 DATA 0,0
8150 DATA 117,117,86,86,28,28,3,3
8155 DATA 0,0,0,0
8160 DATA 41,41,18,18,15,15
8165 DATA 0,0,0,0,0,0
8200 REM: ATLP1,1966
8205 DATA 2378
8210 DATA 21,21,10,10,11,11
```

```
8215 DATA 2,2
8220 DATA 108,108,6,6,17,17,11,11,6,6,5,5,3,3,6,6,37,37,21,21
8225 DATA 5,5
8230 DATA 1,1
8235 DATA 0,0,0,0
8240 DATA 92,92,5,5,10,10,79,79
8245 DATA 0,0
8250 DATA 152,152,115,115,35,35,2,2
8255 DATA 0,0,0,0
8260 DATA 43,45,23,23,18,18
8265 DATA 0,0,0,0,0,0
8300 REM:DENP1,1966
8305 DATA 1681
8310 DATA 9,9,7,7,2,2
8315 DATA 2,2
8320 DATA 105,107,10,10,16,16,8,8,11,11,3,3,9,9,8,8,24,24,19,19
8325 DATA 1,1
8330 DATA 0,0
8335 DATA 0,0,0,0
8340 DATA 73,77,0,0,16,16,57,57
8345 DATA 4,6
8350 DATA 110,110,65,65,40,40,0,0
8355 DATA 0,0,0,0
8360 DATA 40,40,17,17,17,17
8365 DATA 1,1,0,0,1,1
8400 REM: NYTP1,1956
8405 DATA 2810
8410 DATA 27,27,22,22,2,2
8415 DATA 7,7
8420 DATA 131,133,7,7,19,19,18,18,5,5,5,5,13,13,12,12,36,38,27,27
8425 DATA 10,10
8430 DATA 0,0
8435 DATA 0,0,0,0
8440 DATA 145,145,13,13,23,23,109,109
8445 DATA 1,1
8450 DATA 193,193,125,125,68,68,2,2
8455 DATA 1,1,1,1
8460 DATA 80,110,39,39,35,35
8465 DATA 1,1,0,0,1,1
8500 REM:LATP1,1956
8505 DATA 2459
8510 DATA 22,22,17,17,3,3
8515 DATA 2,2
8520 DATA 160,166,2,2,5,5,22,22,4,4,4,4,12,12,11,11,50,50,33,33
8525 DATA 6,6
8530 DATA 0,0
8535 DATA 0,0,0,0
8540 DATA 126,126,1,1,20,20,108,108
8545 DATA 3,3
8550 DATA 173,173,119,119,51,51,4,4
8555 DATA 0,0,0,0
```

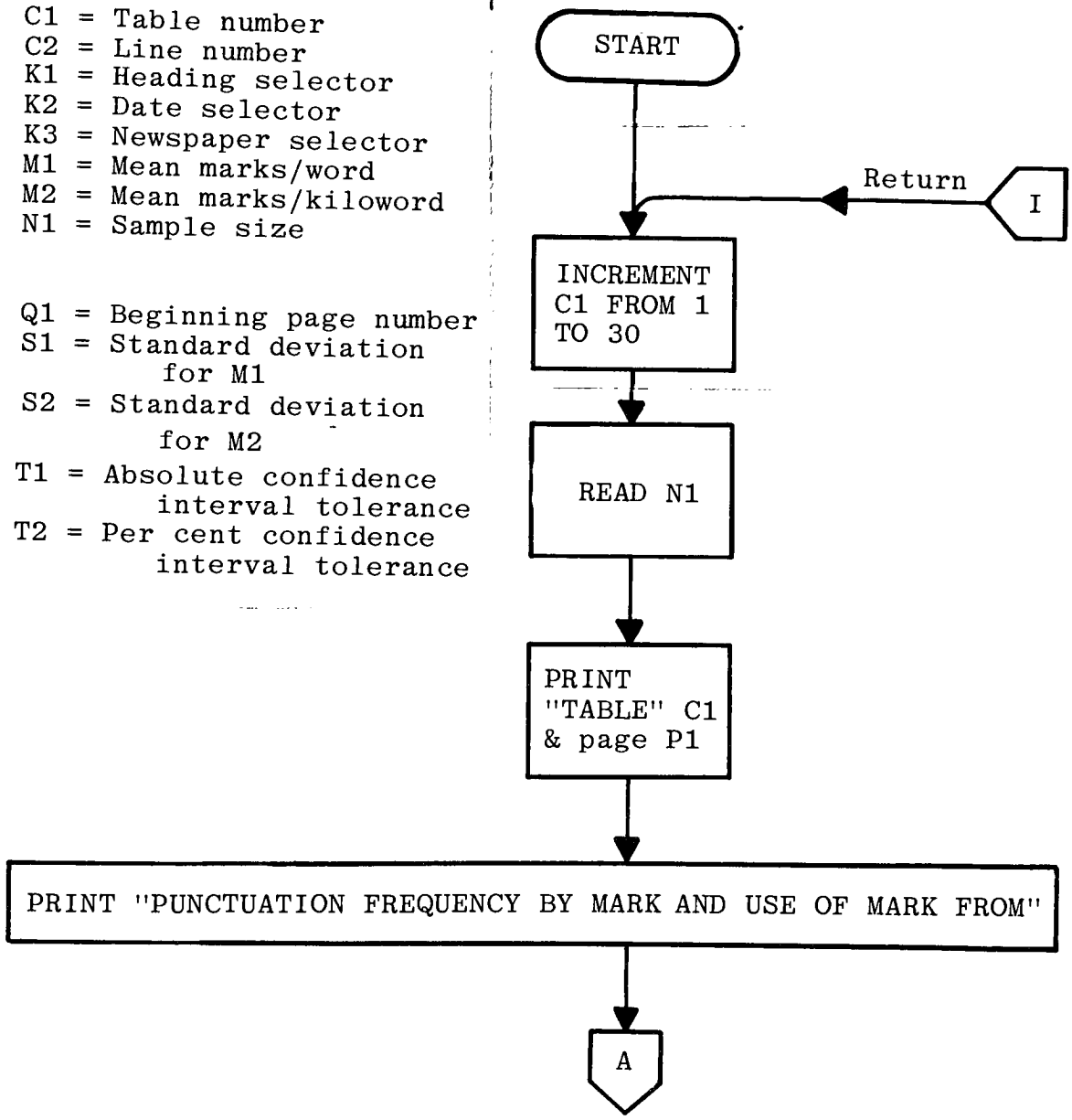
```
8560 DATA 22,26,11,11,11,11
8565 DATA 1,1,1,1,0,0
8600 REM: CHIP1,1956
8605 DATA 2344
8610 DATA 22,22,14,14,8,8
8615 DATA 0,0
8620 DATA 155,159,2,2,7,7,22,22,9,9,3,3,16,16,12,12,54,54,25,35
8625 DATA 5,5
8630 DATA 0,0
8635 DATA 2,2,1,1
8640 DATA 89,91,3,3,10,10,76,76
8645 DATA 6,6
8650 DATA 167,167,113,113,55,55,1,1
8655 DATA 0,0,0,0
8660 DATA 34,34,12,12,12,12
8665 DATA 3,3,0,0,2,2
8700 REM: ATIP1,1956
8705 DATA 3132
8710 DATA 37,37,26,26,9,9
8715 DATA 3,3
8720 DATA 185,187,8,8,12,12,31,31,5,5,6,6,12,12,10,10,55,55,35,35
8725 DATA 7,7
8730 DATA 2,2
8735 DATA 0,0,0,0
8740 DATA 130,130,4,4,16,16,107,107
8745 DATA 8,8
8750 DATA 220,220,160,160,63,63,0,0
8755 DATA 3,3,3,3
8760 DATA 53,57,24,24,23,23
8765 DATA 1,1,0,0,0,0
8800 REM:DENP1,1956
8805 DATA 2445
8810 DATA 20,20,12,12,8,8
8815 DATA 2,2
8820 DATA 122,130,0,0,5,5,18,18,4,4,4,4,6,6,5,5,41,41,22,22
8825 DATA 11,11
8830 DATA 0,0
8835 DATA 0,0,0,0
8840 DATA 72,72,2,2,23,23,49,49
8845 DATA 4,4
8850 DATA 128,128,98,98,29,29,1,1
8855 DATA 1,1,0,0
8860 DATA 4,6,2,2,2,2
8865 DATA 2,2,0,0,0,0
8900 REM: NYT,P1,1946
8905 DATA 3949
8910 DATA 33,33,30,30,0,0
8915 DATA 16,16
8920 DATA 208,234,1,1,13,13,22,22,15,15,7,7,16,16,11,11,58,58,33,33
8925 DATA 3,3
8930 DATA 0,0
```

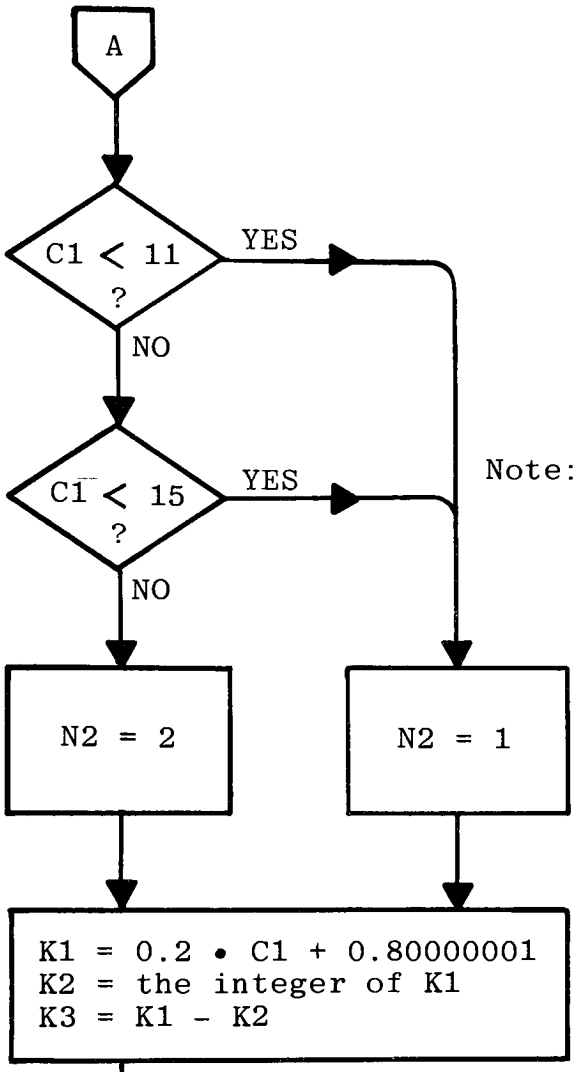
```
8935 DATA 0,0,0,0
8940 DATA 188,190,6,6,20,20,162,162
8945 DATA 15,15
8950 DATA 206,206,149,149,43,43,0,0
8955 DATA 0,0,0,0
8960 DATA 55,65,23,23,22,22
8965 DATA 1,1,0,0,0,0
9000 REM: LAT,P1,1946
9005 DATA 3210
9010 DATA 22,22,20,20,2,2
9015 DATA 3,3
9020 DATA 160,160,8,8,15,15,12,12,5,5,3,3,11,11,5,5,60,60,43,43
9025 DATA 9,9
9030 DATA 0,0
9035 DATA 0,0,0,0
9040 DATA 209,211,6,6,36,36,162,162
9045 DATA 4,4
9050 DATA 188,188,120,120,68,68,0,0
9055 DATA 0,0,0,0
9060 DATA 51,57,25,25,24,24
9065 DATA 1,1,0,0,1,1
9100 REM: CHI,P1,1946
9105 DATA 3614
9110 DATA 41,41,37,37,2,2
9115 DATA 5,5
9120 DATA 218,222,7,7,16,18,39,39,9,9,6,6,16,16,16,16,61,61,41,41
9125 DATA 7,7
9130 DATA 0,0
9135 DATA 0,0,0,0
9140 DATA 121,121,3,3,4,4,113,113
9145 DATA 4,4
9150 DATA 198,198,145,145,56,56,3,3
9155 DATA 0,0,0,0
9160 DATA 75,95,31,31,27,27
9165 DATA 2,2,1,1,1,1
9200 REM: ATL,P1,1946
9205 DATA 3131
9210 DATA 41,41,26,26,12,12
9215 DATA 1,1
9220 DATA 181,181,9,9,7,7,18,18,8,8,11,11,16,16,12,12,62,62,42,42
9225 DATA 6,6
9230 DATA 2,2
9235 DATA 0,0,0,0
9240 DATA 128,128,3,3,30,30,96,96
9245 DATA 2,2
9250 DATA 175,175,128,128,44,44,3,3
9255 DATA 0,0,0,0
9260 DATA 70,80,32,32,30,30
9265 DATA 4,4,4,4,0,0
9300 REM: DEN,P1,1946
9305 DATA 3607
```

```
9310 DATA 30,30,23,23,6,6
9315 DATA 8,8
9320 DATA 253,255,5,5,17,17,70,70,17,17,2,2,20,20,17,17,57,57,43,43
9325 DATA 17,17
9330 DATA 0,0
9335 DATA 1,1,0,0
9340 DATA 141,141,8,8,44,44,89,89
9345 DATA 7,7
9350 DATA 224,224,137,137,82,82,5,5
9355 DATA 2,2,1,1
9360 DATA 60,60,25,25,27,27
9365 DATA 7,7,4,4,3,3
9400 STOP
9401 END
```

C1 = Table number
C2 = Line number
K1 = Heading selector
K2 = Date selector
K3 = Newspaper selector
M1 = Mean marks/word
M2 = Mean marks/kiloword
N1 = Sample size

Q1 = Beginning page number
S1 = Standard deviation for M1
S2 = Standard deviation for M2
T1 = Absolute confidence interval tolerance
T2 = Per cent confidence interval tolerance

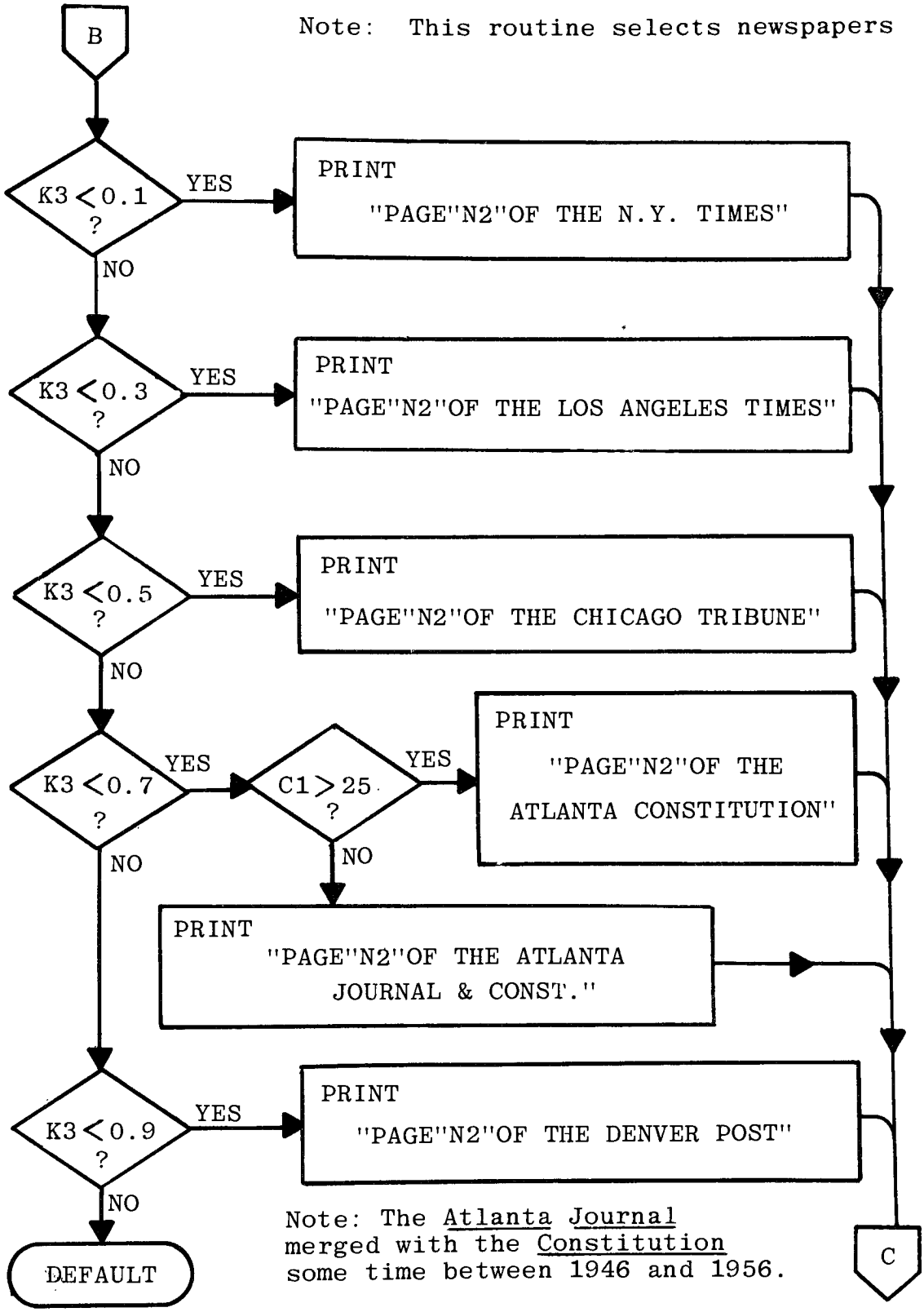


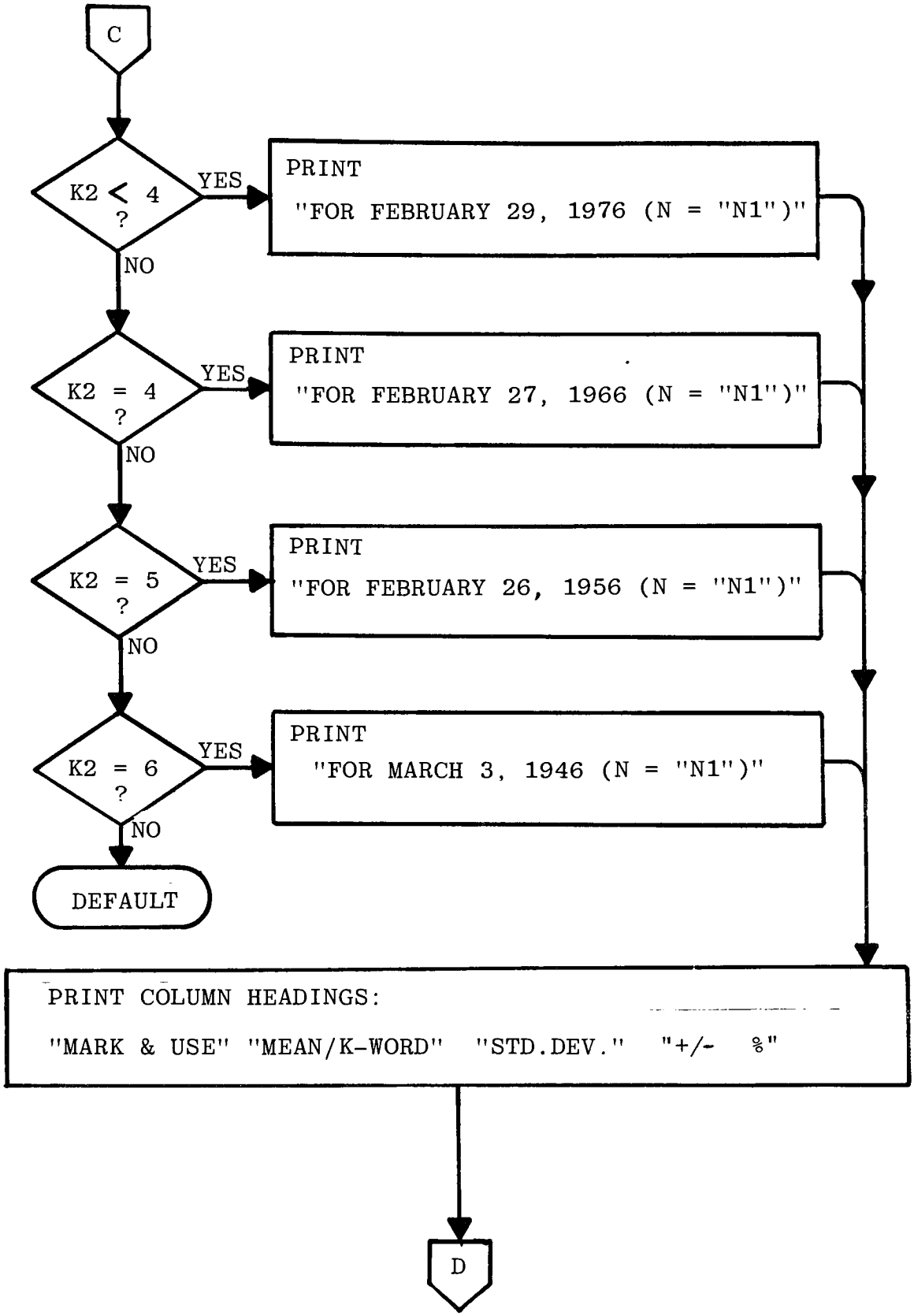


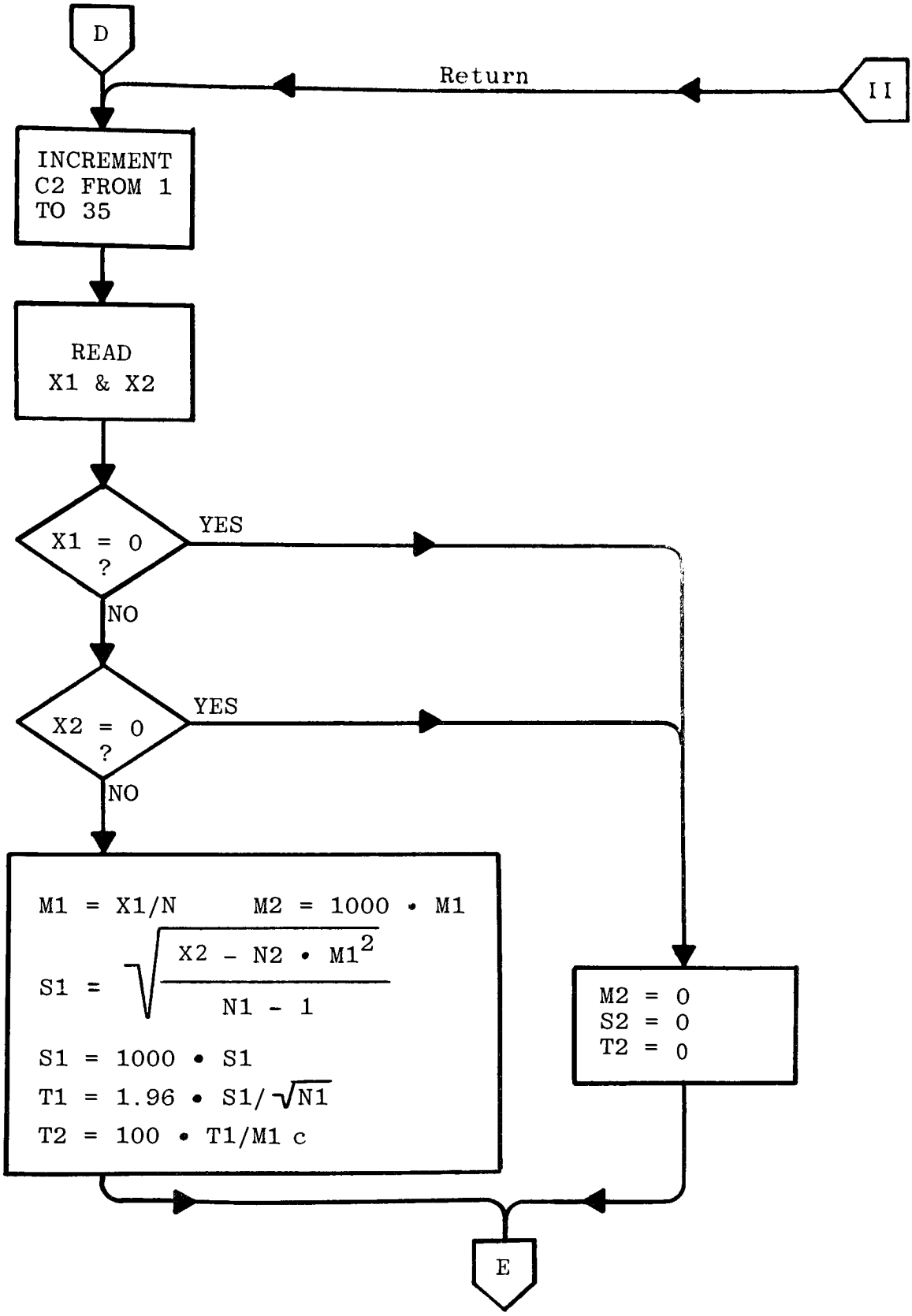
Note: This routine selects the sample page number

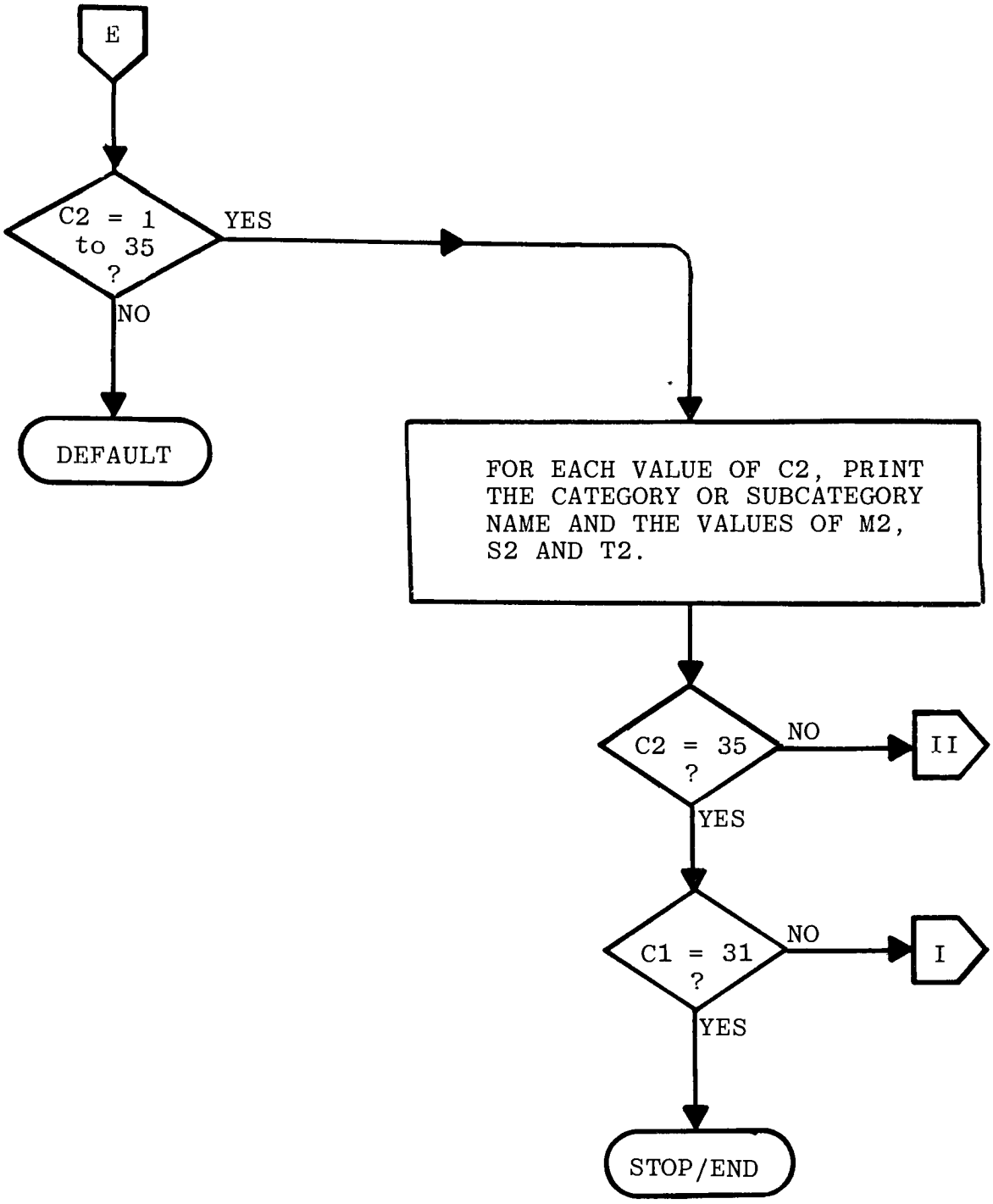
Note: This routine establishes the heading date and newspaper name variables in such a manner that:

K2	Date	K3	Newspaper
1--3	Feb. 29, 1976	0.0	N.Y. Times
4	Feb. 27, 1966	0.2	L.A. Times
5	Feb. 26, 1956	0.4	Chicago Trib.
6	Mar. 3, 1946	0.6	Atlanta J & C
		0.8	Denver Post


B
Note: This routine selects newspapers
K3 < 0.1 ?
YES
PRINT "PAGE"N2"OF THE N.Y. TIMES"
NO
K3 < 0.3 ?
YES
PRINT "PAGE"N2"OF THE LOS ANGELES TIMES"
NO
K3 < 0.5 ?
YES
PRINT "PAGE"N2"OF THE CHICAGO TRIBUNE"
NO
K3 < 0.7 ?
YES
C1 > 25 ?
YES
PRINT "PAGE"N2"OF THE ATLANTA CONSTITUTION"
NO
PRINT "PAGE"N2"OF THE ATLANTA JOURNAL & CONST."
NO
K3 < 0.9 ?
YES
PRINT "PAGE"N2"OF THE DENVER POST"
NO
DEFAULT
Note: The Atlanta Journal merged with the Constitution some time between 1946 and 1956.
C

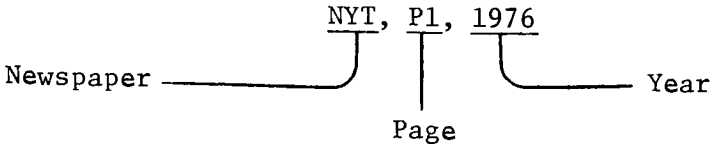

C
K2 < 4 ?
YES
PRINT
"FOR FEBRUARY 29, 1976 (N = "N1")"
NO
K2 = 4 ?
YES
PRINT
"FOR FEBRUARY 27, 1966 (N = "N1")"
NO
K2 = 5 ?
YES
PRINT
"FOR FEBRUARY 26, 1956 (N = "N1")"
NO
K2 = 6 ?
YES
PRINT
"FOR MARCH 3, 1946 (N = "N1")"
NO
DEFAULT
PRINT COLUMN HEADINGS:
"MARK & USE" "MEAN/K-WORD" "STD.DEV." "+/- %"
D


D
Return
II
INCREMENT C2 FROM 1 TO 35
READ X1 & X2
X1 = 0 ?
YES
NO
X2 = 0 ?
YES
NO
M1 = X1/N
M2 = 1000 • M1
S1 = √((X2 - N2 • M1^2)/(N1 - 1))
S1 = 1000 • S1
T1 = 1.96 • S1/√N1
T2 = 100 • T1/M1 c
M2 = 0
S2 = 0
T2 = 0
E


E
C2 = 1 to 35 ?
YES
NO
DEFAULT
FOR EACH VALUE OF C2, PRINT THE CATEGORY OR SUBCATEGORY NAME AND THE VALUES OF M2, S2 AND T2.
C2 = 35 ?
NO
II
YES
C1 = 31 ?
NO
I
YES
STOP/END

NOTES

1. The statistical abstract for Tables D- through D-30 is included in the preceeding computer program as data statements. Data for tables are identified as follows:



2. The statistical abstract given by Tables E-1, E-2, and E-3 include the data upon which Tables VI through XII are based. In these tables, the variables are defined as follows:

n = the number of words in a given cell of the study.

ΣX = the sum of all the occurrences of marks of punctuation within cells.

ΣX^2 = the sum of each occurrence squared.

$\bar{X}$ = the mean number of marks of punctuation per word.

s = the standard deviation of the mean.

3. In both the preceeding and following statistical abstracts, the following abbreviations are used:

NYT = The New York Times — 1976 = February 29, 1976

LAT = The Los Angeles Times — 1966 = February 27, 1966

CHI = The Chicago Tribune — 1956 = February 26, 1956

ATL = The Atlanta Journal & Constitution (except 1946, Atlanta Constitution — 1946 = March 3, 1946

DEN = The Denver Post

TABLE E-1

STATISTICAL ABSTRACT: ALL MARKS OF PUNCTUATION

YEAR	PAPER	n	ΣX	ΣX^2	$\bar{X}$	s
1976	NYT	2135	449	554	0.2030	0.4640
	LAT	2201	424	492	0.1926	0.4319
	CHI	2107	460	558	0.2183	0.4661
	ATL	1649	319	385	0.4429	0.4429
	DEN	2671	521	651	0.1951	0.4536
	ALL	10,763	2173	2640	0.2019	0.4523
1966	NYT	2477	502	616	0.2112	0.4633
	LAT	2273	428	562	0.1884	0.4694
	CHI	2011	405	501	0.2014	0.4568
	ATL	2378	427	519	0.1796	0.4314
	DEN	1681	342	422	0.2035	0.4580
	ALL	10,720	2104	2620	0.1963	0.4538
1956	NYT	2810	595	716	0.2117	0.4583
	LAT	2459	520	632	0.2115	0.4608
	CHI	2344	489	605	0.2086	0.4633
	ATL	3132	669	833	0.2136	0.4695
	DEN	2445	373	443	0.1526	0.3975
	ALL	13,190	2646	3229	0.2006	0.4523
1946	NYT	3949	741	919	0.1876	0.4445
	LAT	3210	655	781	0.2040	0.4491
	CHI	3614	679	822	0.1879	0.4384
	ATL	3131	612	751	0.1955	0.4491
	DEN	3607	750	876	0.2079	0.4469
	ALL	17,511	3437	4149	0.1963	0.4454
ALL	NYT	11,271	2278	2805	0.2021	0.4561
	LAT	10,143	2027	2467	0.1998	0.4827
	CHI	10,076	2033	2486	0.2018	0.4539
	ALT	10,290	2027	2488	0.1970	0.4506
	DEN	10,404	1986	2392	0.1909	0.4398
ALL	BIG	31,490	6347	7758	0.2016	0.4536
	SMALL	20,694	4013	4880	0.1939	0.4452
GRAND SUMMARIES		52,184	10,360	12,638	0.1985	0.4503

TABLE E-2

STATISTICAL ABSTRACT: TERMINAL MARKS OF PUNCTUATION

YEAR	PAPER	n	$\Sigma X, \Sigma X^2$	$1/\bar{X}$	$1/s$
1976	NYT	2135	68	31.39	5.693
	LAT	2201	78	28.22	5.408
	CHI	2107	103	20.46	4.637
	ATL	1649	88	18.74	4.448
	DEN	2671	108	24.73	5.076
	ALL	10,763	445	24.19	5.023
1966	NYT	2377	94	25.29	5.130
	LAT	2273	95	23.93	4.996
	CHI	2011	89	22.60	4.861
	ATL	2378	115	20.68	4.660
	DEN	1681	65	25.86	5.185
	ALL	10,720	458	23.41	4.945
1956	NYT	2810	126	22.30	4.831
	LAT	2459	119	20.66	4.658
	CHI	2344	114	20.56	4.648
	ATL	3132	163	19.21	4.501
	DEN	2445	98	24.95	5.971
	ALL	13,190	620	21.27	4.725
1946	NYT	3949	149	26.50	5.247
	LAT	3210	120	26.75	5.271
	CHI	3614	145	24.92	5.095
	ATL	3131	128	24.46	5.049
	DEN	3607	138	26.14	5.212
	ALL	17,511	680	25.75	5.176
ALL	NYT	11,271	437	25.79	5.180
	LAT	10,143	412	24.61	5.065
	CHI	10,076	451	22.34	4.836
	ATL	10,290	494	20.83	4.677
	DEN	10,404	409	25.44	5.145
ALL	BIG	31,490	1300	24.22	5.992
	SMALL	20,694	903	22.92	4.895
GRAND SUMMARIES		52,184	2203	23.69	4.973

TABLE E-3

STATISTICAL ABSTRACT: INTRASENTENCE MARKS OF PUNCTUATION

YEAR	PAPER	n	ΣX	ΣX^2	$\bar{X}$	s
1976	NYT	2135	381	486	0.1784	0.4426
	LAT	2201	346	414	0.1572	0.4043
	CHI	2107	357	455	0.1694	0.4328
	ATL	1649	231	297	0.1401	0.4007
	DEN	2671	413	543	0.1546	0.4236
	ALL	10,763	1728	2195	0.1606	0.4221
1966	NYT	2377	408	522	0.1716	0.4361
	LAT	2273	333	467	0.1465	0.4290
	CHI	2011	316	416	0.1571	0.4025
	ATL	2378	312	404	0.1312	0.4245
	DEN	1681	277	357	0.1648	0.4305
	ALL	10,720	1646	2162	0.1535	0.4220
1956	NTY	2810	469	590	0.1669	0.4268
	LAT	2459	401	513	0.1711	0.4355
	CHI	2344	375	491	0.1600	0.4289
	ATL	3132	506	670	0.1616	0.4335
	DEN	2445	275	345	0.1125	0.3585
	ALL	13,190	2026	2609	0.1536	0.4174
1946	NTY	3949	592	770	0.1499	0.4154
	LAT	3210	535	661	0.1667	0.4221
	CHI	3614	534	677	0.1664	0.4281
	ATL	3131	484	623	0.1546	0.4185
	DEN	3607	612	738	0.1698	0.4194
	ALL	17,511	2757	3469	0.1574	0.4163
ALL	NTY	11,271	1850	2368	0.1641	0.4280
	LAT	10,143	1615	2055	0.1592	0.4210
	CHI	10,076	1582	2035	0.1570	0.4211
	ATL	10,290	1533	2003	0.1490	0.4153
	DEN	10,404	1577	1983	0.1516	0.4094
ALL	BIG	31,490	5047	6458	0.1603	0.4236
	SMALL	20,694	3110	3877	0.1503	0.4118
GRAND SUMMARIES		52,184	8157	10,435	0.1563	0.4190

APPENDIX F

CURRICULUM FORMULAE

CURRICULUM FORMULAE

For the reasons explained in Chapter IV, the data was collected in the form of sums of the occurrences (ΣX) and sums of the occurrences squared (ΣX^2). Programs for calculating statistics typically require input of the occurrences (e.g., Statistical Programs for the Social Sciences, and many hand-held calculators). These programs would require a huge number of data entries to be made including large numbers of zeros. The average cell would have required 2609 entries of which 2091 would be zeros, and to repeat this 35 times for each of 20 cells for a total of 1,826,300 entries. Of these, 1,463,700 would have been zeros. Obviously a method to avoid this would be desirable. One method of avoiding this is to develop a program which allows the needed calculations to be made from the sums of occurrences and the sums of the occurrences squared. Such a program was developed and is shown written in BASIC in Appendix E.

For all calculations made in this study, the sum of the occurrences and the sum of the occurrences squared were used as the basic data source. These data were used for calculating the mean occurrences, the standard deviations of the means, the confidence interval tolerances, and in analyses of variance. This appendix traces the calculations from ΣX and ΣX^2 to the final reported statistics.

Data Summation

The number of occurrences of each mark of punctuation (ΣX) and the squares of each occurrence (ΣX^2) were summed for the twelve punctuation mark categories and the 23 subcategories for each of the twenty cells in this study. These categories and subcategories are identified in Appendix C and the process of collecting these data is fully explained in Chapter IV.

Calculation of the Mean, Standard Deviation

The mean and standard deviation were calculated using formulae (1) and (2).

$$\bar{X} = \Sigma X/n \qquad (1)^1$$

where: $\bar{X}$ = the mean (punctuation frequency)

ΣX = the sum of the occurrences

n = number of words in the sample

$$s = \sqrt{\frac{\Sigma X^2 - n \cdot \bar{X}^2}{n - 1}} \qquad (2)^2$$

where: s = the standard deviation

ΣX^2 = the sum of the occurrences squared

n = number of words in the sample

$\bar{X}$ = the mean, q.v. formula (1)

Calculation of Confidence Intervals

The confidence intervals were calculated using formula (5). First, using the method described by Rickmers (1967), confidence intervals were calculated.

$$\bar{X} - \frac{t_{\alpha/2}\,\sigma}{n} < \mu < \bar{X} + \frac{t_{\alpha/2}\,\sigma}{n} \qquad (3)^3$$

where: $\bar{X}$ = the mean (punctuation frequency)

$t_{\alpha/2}$ = tabular value of student-t for a given value of .

σ = the standard deviation

n = the number or words in the sample

Formula (3) states that--based upon a sample size of n and a mean of $\bar{X}$ --the true mean (μ) will fall between the two limits given, at a given probability level. For all calculations in this study, 95% probability level was used, or conversely stated, $\alpha = 0.05$. Since the value of n is large in all cases in this study (never less than 300), we find the value of $t_{\alpha/2}$ in all cases from Table A.2 of Rickmer's text[4] to be 1.96. The formula (3) may now be restated as formula (4).

$$T_1 = \pm \frac{1.96 \times s}{\sqrt{n}} \qquad (4)$$

where: T_1 = the confidence intervals on the mean

s = the standard deviation of the sample mean

n = the number of words in the sample

It is useful to state this $\pm$ value as a percentage of the mean for better understanding of the quality of the means. To accomplish this, each confidence interval was divided by the mean and multiplied by 100.

$$T_2 = \pm \frac{196 \times s}{\bar{X} \times \sqrt{n}} \qquad (5)$$

where: T_2 = the $\pm$ tolerance of a mean

s = the standard deviation of a mean

$\bar{X}$ = the mean (punctuation frequency)

n = the number of words in the sample

Hypothesis Test of Means Between Page 1 and Page 2 Samples

In Section IV-1.2 of Chapter IV, hypothesis tests were introduced to determine if significant differences exist between samples taken from pages 1 and pages 2 of the five chosen newspapers for 1976 (q.v. Table IV). Means and standard deviations were calculated for page 1 and page 2 of each of the five newspapers, i.e., ten means and ten standard deviations. Next, the degrees of freedom were found for each using formula (6).

$$\nu = n - 1 \qquad (6)^4$$

Where: ν = degrees of freedom

n = number of words in the sample

From the standard deviations and degrees of freedom for each of the ten cells in these hypothesis tests, the value of the pooled standard deviation across pages 1 and 2 for each of the five newspapers was found using formula (7).

$$S_p = \frac{\nu_1 \quad S_1^2 + \nu_2 \, S_2^2}{\nu_1 + \nu_2} \qquad (7)^5$$

where: S_p = pooled standard deviation

ν = degrees of freedom for pages $_1$ and $_2$

s = standard deviation for pages $_1$ and $_2$

Next, the t statistic was calculated using formula (8).

$$t_{(\nu_1 + \nu_2)} = \frac{\bar{X}_1 - \bar{X}_2}{S_p \sqrt{1/n_1 + 1/n_2}} \qquad (8)^6$$

where: $t_{(\nu_1 + \nu_2)}$ = t statistic

$\bar{X}$ = mean for pages $_1$ and $_2$

n number of words in the sample

S_p = pooled standard deviation

Calculation of Words Per Sentence

The number of words per sentence is shown in two places in this study (q.v., Table VI and Table IX). In both cases, the frequency of terminal marks of punctuation were calculated, and then the reciprocal of these numbers were found.

Analyses of Variance Between Cells

In Chapter VII, six analyses of variance were reported. The 'sum of squares' method described by Rickmers[7] was used. Statistics presented in Tables XI, XII, and XIII were calculated using the Texas Instruments model SR-52 with statistical program ST1-05. First, the total sum of squares was calculated using formula (9).

$$SST = \sum X_{ij}^2 - \frac{(T..)^2}{n} \qquad (9)^7$$

where: SST = sum of the squares, total

$\sum X_{ij}$ = the sum of all occurrences

T.. = the sum of the square of each occurrence

the sum of squares for the effect of interest (either time or region) was calculated using formula (10).

$$SSR = \sum \frac{(T_{i.})^2}{nj} - \frac{(T..)^2}{n} \qquad (10)^8$$

where: SSR = the sum of squares for the effect of interest

T_i = the sum of the squares for the effect of interest

T.. = the sum of the squares of each occurrence

nj = number of words in the cell for the effect of interest

n = the total number of words in the experiment

Fourth, the residual (or error) was calculated using formula (11).

$$SSE = \sum X_{ij}^2 - \frac{T_{i.}^2}{j} \qquad (11)^9$$

where: SSE = the sum of the squares, residual

$\sum X_{ij}$ = the sum of all occurrences

$T_{i.}$ = the sum of the squares of each occurrence not included in the effect of interest.

Fifth, a check was made to insure that, in fact equality existed as stated by formula (12).

$$SSE = SST - SSR \qquad (12)^{10}$$

where: SSE = the sum of the squares, residual

SST = the sum of the squares, total

SSR = the sum of the squares of the effect of interest

Sixth, the mean squares were found by dividing the sum of the squares by the degrees of freedom (q.v., Tables XI, XII, and XIII).

Seventh, the F statistics were calculated using formula (13).

$$F = s_1^2/s_2^2 \qquad (13)^{11}$$

where: F = the F statistic

s^2 = variance for effects $_1$ and $_2$

Last, the values for F were taken from the F table[12] for given degrees of freedom for each effect and for a 95% probability level (α = 0.05). These values will be found listed in Tables XI, XII, and XIII.

The actual performance of these calculations was done using the Texas Instruments model SR-52 using program ST1-05. Because this program is designed for direct entry of the data, and because of the nature of the data collection methods (q.v., beginning of Curriculum Formulae), the memories of the calculator were accessed directly and entries made in them. These calculations were done with the assistance of Dr. Kathleen E. Crandall.

ENDNOTES, APPENDIX F

[1]Albert D. Rickmers, and Hollis N. Todd, Statistics: An Introduction, New York: McGraw-Hill Book Co., 1967, p. 19

[2]Lawrence L. Lapin, Statistics for Modern Business Decisions, New York: Harcourt Brace Jovanovich, Inc., 1973, p. 57.

[3]op. cit., p. 94.

[4]ibid, pp. 79--81.

[5]ibid, p. 85.

[6]ibid, p. 85.

[7]ibid, p. 162.

[8]ibid, p. 163.

[9]ibid, p. 163.

[10]ibid, p. 163.

APPENDIX G

GLOSSYNOGRAPHIC WRITING LABORATORY

THE GLOSSYNOGRAPHIC WRITING LABORATORY

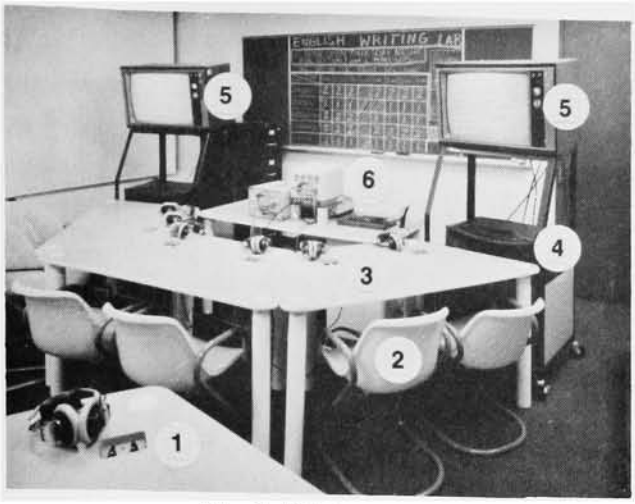


Fig. 1. Writing Laboratory

Fig. 2. Teacher's Control Station

Figure 1 shows the writing laboratory where students sit (2) and practice writing correct English each week. Students with sufficient hearing to profit from acoustic stimuli use headsets (1) with individual binaural controls to listen to prerecorded lessons. Words appear on the student television screens (5) in synchronism with the sound. A video playback unit (4) is also available for other applications. The teacher's control station (6) is shown in detail in Figure 2.

Figure 2 shows the teacher's control station in the writing laboratory. Messages may be typed on the keyboard (1), or reproduced from the word recorder/reproducer (2). Teachers may speak using the microphone (4), or prepared lessons may be played on the tape recorder/reproducer (5), and amplified by the sound amplifier (6). All equipment power is controlled by the switch (7). A slide projector (not shown) displays coordinated pictures automatically at the appropriate time in the lesson.

Figure 3 shows a block diagram of the glossynograph system. The *sound recorder/reproducer* provides an audio signal to the *sound amplifier*. The amplified sound is distributed through *individual binaural controls* to twelve student *headsets*. The *sound recorder/reproducer* also provides a synchronizing pulse at the end of most paragraphs. This pulse is sent through the junction box to the *print recorder/reproducer and* starts it 'reading' words. The words —in digital form—are sent to the keyboard where they are changed into a video signal. The video signal is then displayed by the *teacher's* and *students' television monitors* at the same time.

The *print recorder/reproducer* is programmed so that it automatically stops at the end of most paragraphs. When the next synchronizing pulse is received from the *sound recorder/reproducer*, it begins 'reading' the words again. In that way, the words are automatically resynchronized at the beginning of most paragraphs.

Pictures are automatically coordinated with print. Picture control pulses are recorded on the *print recorder/reproducer* in coordination with the words so that an appropriate picture may be displayed with the words in the lessons. These picture control pulses are reproduced by the *print recorder/reproducer* and pass through the *junction box* to the picture projector and cause it to display either the next picture or the last picture in the projector tray.

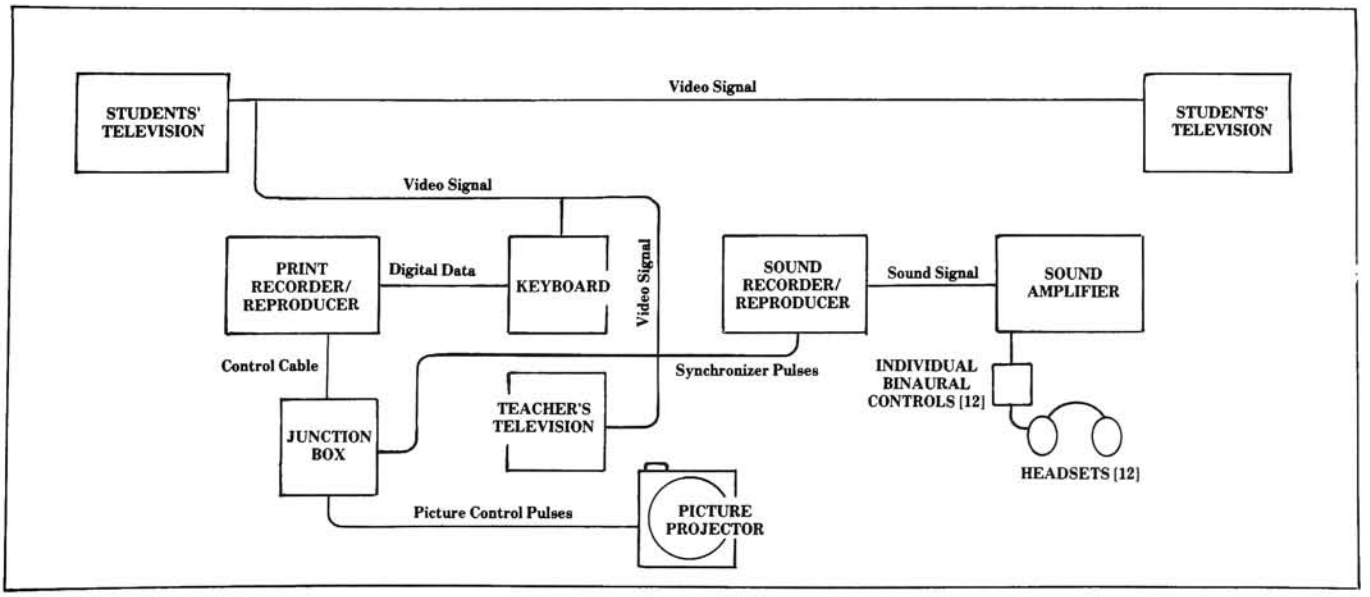


Fig. 3. System Block Diagram